Combined brain imaging methods to study structural and functional signatures of auditory and speech processing

Thesis (cumulative thesis)
Presented to the Faculty of Arts and Social Sciences
of the University of Zurich
for the Degree of Doctor of Philosophy

by

Franziskus Liem

Accepted in the Autumn Term 2012
on the Recommendation of the Doctoral Committee:

Prof. Dr. Martin Meyer (main advisor)
Prof. Dr. Volker Dellwo

Zurich, 2013

Contents

Acknowledgements

I would like to express my deepest gratitude to Martin Meyer for fostering me when I was little and for his tremendous financial and scientific support over the years. Thank you for your guidance and for providing an environment in which I could follow my bizarre interests.

I am also extremely grateful to Lutz Jäncke for welcoming me at his department and for his scientific and financial support.

Thanks to Volker Dellwo for being part of the committee.

I also would like to thank Kai Lutz, Roger Lüchinger, Martina Hurschler, Tino Zaehle and Anja Burkhard for their cooperation on my projects.

This work would not have been possible without the volunteers that patiently endured our experiments. Thank you for that.

I also would like to thank all my colleagues at the Department Neuropsychology and the INAPIC for the fun time. Special thanks to Yolanda for sharing an office, to Ladina for providing the LaTeX template, and to Jürgen and Nicki for sparking my interest in anatomical and network analyses.

Most of all, I would like to thank my family, friends, and Nicole.

Abstract

No clear consensus has been reached regarding the brain organization that underlies language and speech functions, despite the long history of research in this field. In three empirical studies, this thesis addresses several open questions in the field of cognitive auditory and speech neuroscience and emphasizes the importance of using a combination of several neuroimaging approaches to target scientific problems from several perspectives.

Study I was conducted to optimize the timing of auditory functional magnetic resonance imaging (fMRI) setups. Findings from this study allow to design substantially improved auditory fMRI experiments.

Study II tested the "asymmetric sampling in time" (AST) framework. This was done by implementing an established behavioral experiment in a clustered-sparse fMRI study. The results clearly support AST's predictions with regards to the functional lateralization in primary and non-primary auditory cortex. Additionally, novel insight into structural lateralization and auditory cognition is presented.

In Study III the relationship between the electroencephalogram's (EEG) auditory evoked potential's (AEP) component N1 and macroanatomical characteristics of the supratemporal plane was examined. Cortical thickness was measured via surface-based morphometry. The results show that cortical thickness in the supratemporal plane predicts the amplitude of the N1 modulation. This electrophysiological-structural mapping approach might allow future EEG studies to localize the brain sources of scalp activity with an extremely high spatial precision, in the millimeter range.

In closing, I will discuss possible cellular mechanisms that underlie individual differences and change in macroanatomy and will sketch future directions in which neuroscientific progress can be expected.

Zusammenfassung

Obwohl es ein breites Spektrum an neurowissenschaftlichen Untersuchungen zu den neuronalen Grundlagen der Kognition von Sprach- und Lautsprachfunktionen gibt, hat sich kein breiter Konsens in diesem Feld gebildet. Diese Arbeit untersucht mittels drei empirischer Studien einige offene Fragen zu auditorischen und sprachbezogenen Themen auf dem Gebiet der Kognitiven Neurowissenschaft. Ein Fokus wird auf die Kombination verschiedener Bildgebungsmethoden gelegt um wissenschaftliche Fragestellungen von verschiedenen Seiten zu beleuchten.

In Studie I wurden zeitliche Parameter des Darbietungsdesigns für auditorische funktionelle Magnet Resonanz Tomographie (fMRT) Studien optimiert. Die Ergebnisse dieser Studie erlauben es, deutlich verbesserte auditorische fMRT Studien durchzuführen.

Studie II testete die „Asymmetric Sampling in Time“ (AST) Hypothese in einem clustered-sparse fMRT Design mithilfe eines etablierten Verhaltensexperiments. Die Resultate untermauern die Vorhersagen der AST Hypothese bezüglich der funktionellen Lateralisierung von Regionen des primären und nicht-primären auditorischen Kortex'. Zusätzlich werden auch neue Erkenntnisse über strukturelle Asymmetrie und auditorische Kognition berichtet.

In Studie III wurden Zusammenhänge zwischen dem Elektroenzephalogramm (EEG) und makroanatomischen Eigenschaften des supratemporalen Planums untersucht. Die N1 Komponente des Auditorisch Evozierten Potentials (AEP) wurde mit Massen der kortikalen Dicke korreliert. Dieses strukturelle Mass wurden mithilfe von Oberflächenbasierter Morphometrie gewonnen. Die Studie zeigte, dass kortikale Dicke im supratemporalen Planum die N1 Amplitude vorhersagt.

Zum Abschluss werden mögliche zelluläre Mechanismen diskutiert, die den interindividuellen Unterschieden in der Makroanatomie zugrunde liegen könnten. Des weiteren werden mögliche neurowissenschaftliche Forschungsfelder und Stossrichtungen aufgezeigt, in denen in der Zukunft wichtige wissenschaftliche Erkenntnisse erwartet werden können.

1. Introduction

Presently, scientists have remarkable methods at their disposal to study the brains of healthy human beings in a noninvasive manner. Via positron emission tomography (PET), magnetic resonance imaging (MRI) and electrophysiological techniques, functional and structural brain maps can be acquired and related to cognition. In the present thesis, I will sketch the development of these neuroimaging methods, which led to the emergence of the field of cognitive neuroscience (chapter 1.1). I then will focus on the progress that has been achieved in neuroimaging with regards to the examination of auditory and speech functions, will identify conceptual dead ends that might have resulted in stagnancy in the field and will present alternative approaches (chapter 1.2). Lastly, I will present advanced neuroimaging methods designed specifically to study auditory and speech functions in the brain (chapter 1.3). Emphasis was put on the combination of several neuroimaging methods to consider a scientific problem from different sides. Up until recently, the majority of neuroimaging studies made use of only one of those methods at a time. However, as those methods cover different aspects of brain function and structure, combining them into multimodal imaging settings will be a decisive step towards scientific progress (Groves et al., 2012). This thesis provides examples as how a combination of methods might be used to consider different aspects of a scientific problem to end up with a multi-facet framework.

Through the introduction, three open scientific questions will be considered that will be addressed via three empirical studies. One examination (Study II, chapter 2.2) will present an innovative approach, using functional and structural MRI to examine functional and structural asymmetries in the posterior temporal cortex in the context of a novel speech processing paradigm. The remaining two studies focus on methodological issues. Study I (chapter 2.1) aims to optimize auditory fMRI protocols. Insight from this study is implemented in Study II. Finally, Study III (chapter 2.3) uses a multimodal electrophysiological-structural mapping approach to characterize the basic relationship between electrophysiological signals and macroanatomical brain structure.

1.1. History of human brain imaging and cognitive neuroscience

While neuroimaging methods to study healthy human volunteers have been developed and applied over the last couple of decades, the groundwork that has enabled these methods began to be conducted over a century ago. Functional imaging meth-

ods like PET and fMRI rely on the fact that changes in local neuronal activity results in changes in blood flow and metabolism. Impressive work observing the relationship between cognition and blood flow was published as early as 1881 (Mosso, 1881). The author studied patients with a defect in their skull, which allowed Mosso to measure their brains' pulsation (cf. Figure 1.1). He reported increased pulsation over the prefrontal cortex when patients performed an arithmetic task. Over the century that followed, a plethora of researchers examined the relationship of cognition and blood flow, primarily in lab animals. Insight from these works later resulted in techniques to study brain function in living human beings (Raichle, 2009).

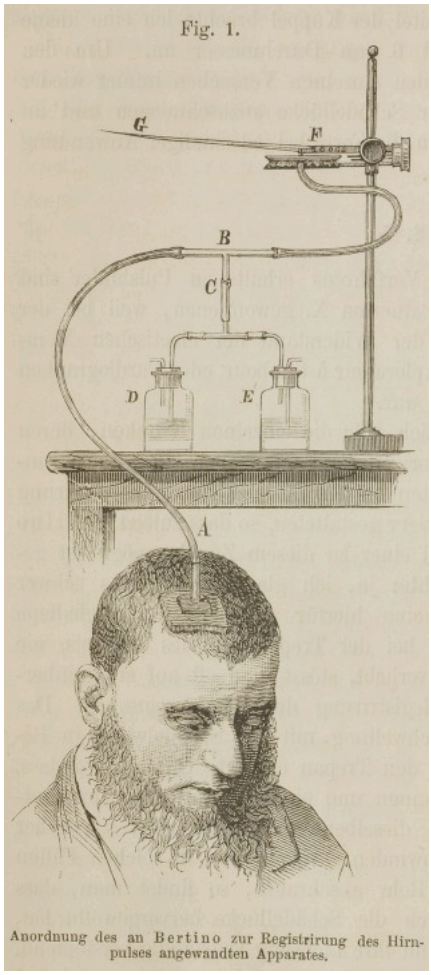

Figure 1.1.: Mosso's apparatus to study local brain pulsation in a living human (Mosso, 1881).

The first neuroimaging method to study brain anatomy in living humans was computer tomography (CT), introduced in 1971 (Raichle, 2009). With this method, which is based on focused X-ray beams, the first three-dimensional images of the brain were acquired. By the late 70s and early 80s, this development inspired the first in-vivo method to study brain function in humans, PET (Raichle, 1983). In its most popular

form, PET monitors regional cerebral blood flow (rCBF). This is achieved by marking molecules in the brain via a radioactive tracer substance and then recording emissions produced by the tracer. The temporal resolution of the image acquisition, which is determined by the tracer's short half-life, could be as low as under one minute, which is far from impressive by today's standards. With this method it became possible to record up to 12 measurements in one subject. While this procedure constituted revolutionary possibilities to study brain function, a critical drawback is the need to inject radioactive tracer into the subject's veins. This makes the procedure unpleasant and limits the possibilities to study the same subject repeatedly over time.

Co-occurring with the development of CT and PET, MRI was introduced. Based on research conducted in the 1940s, 50s and 60s, first structural images of the anatomy of alive humans were recorded in the 1970s (Hinshaw, Bottomley, & Holland, 1977). MRI can measure magnetic properties of tissue in the body. Importantly, different tissue types (for instance the brain's gray and white matter) show different magnetic properties. These different properties can be measured and reconstructed in high-resolution (millimeter range) three-dimensional images of, for instance, neuroanatomy. This application is referred to as structural MRI (sMRI).

Akin to different tissue types, differently oxygenated blood also shows different magnetic properties: while oxygenated hemoglobin is paramagnetic and does not disturb the magnetic field, deoxygenated hemoglobin is diamagnetic, and therefore disturbs a magnetic field. These different properties have long been known. Based on observations by Michael Faraday in 1845, Pauling and Coryell (1936) were the first to measure those different properties (Raichle, 2009). In combination with more recently conducted work (e.g. Thulborn, Waterton, Matthews, & Radda, 1982), this culminated in the discovery of the blood-oxygen-level-dependent (BOLD) signal (Ogawa, Lee, Nayak, & Glynn, 1990). This signal describes the local ratio of oxygenated and deoxygenated hemoglobin and locally changes as a function of neuronal activity. Hence, BOLD functional MRI (fMRI) records an indirect measurement of brain activity (Raichle, 2009).

Since its advent, structural and functional MRI has revolutionized and dominated the field of cognitive neuroscience. Not only is it now possible to study brain structure and function with one imaging technique (PET is not able to acquire structural images), MRI constitutes a noninvasive method. Radioactive tracers or rays are not required to measure structural and functional MR contrasts. As important as the technical advances in acquisition methodology have been, progress in experimental design must not be underestimated.

The first experiments to study human auditory cognition, which were implemented in PET studies, used subtraction designs, originating from the reaction time subtraction method introduced in the 1860s by Franciscus C. Donders (Raichle, 2009). This influential concept in cognitive psychology aims to isolate a cognitive task by measuring and subtracting the reaction times evoked by two different tasks, one experimental and one control task. It is assumed that by subtracting the control reaction time from the

experimental reaction time, only processes specific to the experimental tasks are supposed to be measured. The first human brain mapping experiments that applied this approach presented subjects with, for instance, written words they either had to read out loud (experimental condition) or read silently (control condition). By subtracting the maps acquired during silent reading from the ones acquired during loud reading, scientists aimed to localize the brain regions responsible for processes like articulator code, motor programming, motor output (Petersen, Fox, Posner, Mintun, & Raichle, 1988). The inherently low temporal resolution of PET, in the range of minutes, only allowed to implement simple block designs: in the first block subjects silently read words for one minute, followed by a one minute block of reading out loud. Hence, the simple and straightforward analysis via subtraction methods seemed appropriate at the time.

However, these subtraction designs rely on the principle of "pure insertion", i.e. it assumes that a cognitive process can be added to the task without influencing the, or interacting with the other cognitive processes. For example, if the difference between a "syntax + semantics" task and a "syntax only" task is calculated to isolate semantic processing, the presence of semantic processes must not influence syntactic processing. However, in the case of syntax and semantics this separation cannot be made clearly (Bates & Goodman, 1997). More generally, the assumption of pure insertion is very rarely met (Friston et al., 1996). Since fMRI made it possible to study the brain's reaction to single events or stimuli, as compared to whole blocks of stimuli in PET, more sophisticated design and analysis approaches were possible. These so-called event-related fMRI acquisition designs made it possible to implement factorial and parametric analyses that do not rely on pure insertion. With factorial designs interaction between cognitive processes, violating the assumption of pure insertion, could be modeled explicitly. Parametric designs made it possible to study one task under gradually different conditions, e.g. different working memory load (Friston et al., 1995). With these new designs it became possible to present stimuli in a randomized manner and to consider the subjects' behavioral response to each stimulus. Taken together, these advances in experimental designs and data analysis approaches enabled scientists to examine a wider array of research questions in a more sophisticated manner.

1.2. History of the cognitive neuroscience of language and speech processing

Before the introduction of neuroimaging tools, inferences on the relationship between language/speech and the brain were drawn from patients with aphasia, acquired language and speech impairments due to brain lesions. After the patients' death, their brains were examined and local macro-lesions were connected to deficits in language and speech processing. In fact, finding the brain basis of language and speech processing was one of the first challenges modern neuroscience attended to, in the second half of the

19th century. Pioneering work was conducted by Broca (1861) and Wernicke (1874). The French neurologist Paul Broca related severe speech production deficits to a lesion in the left inferior frontal gyrus, now known as Broca's area. Shortly after Broca, the German neurologist Carl Wernicke related damage of the left posterior superior temporal gyrus, now known as Wernicke's area, to lexical impairments (fluent nonsensical production accompanied by comprehension deficits). Wernicke already appreciated the importance of fibers connecting the temporal and the frontal region. He proposed that sounds of words are stored in Wernicke's area, while motor images of speech are stored in Broca's area. Wernicke's area is hypothesized to be connected to Broca's area in order to transform sound to motor representations when repeating a perceived word. Representations in Wernicke's area are assumed to be of a purely sensory nature and are connected to the memory representations of concepts they describe, in order to give meaning to the sounds (Poeppel & Hickok, 2004). Wernicke's model was later extended by Lichtheim (1885) and was known as the "classical model" of the neuroanatomy of language or Wernicke-Lichtheim-model (cf. Figure 1.2). A key difference between Wernicke's and Lichtheim's model was Lichtheim's proposal of a conceptual center. Wernicke assumed conceptual representations to be spread over the cortex (Eling, 2011). In the decades that followed, the model lost popularity but was brought back into scientific discourse by the work of Norman Geschwind (1965a, 1965b) and is still the prevailing model in the clinical context.

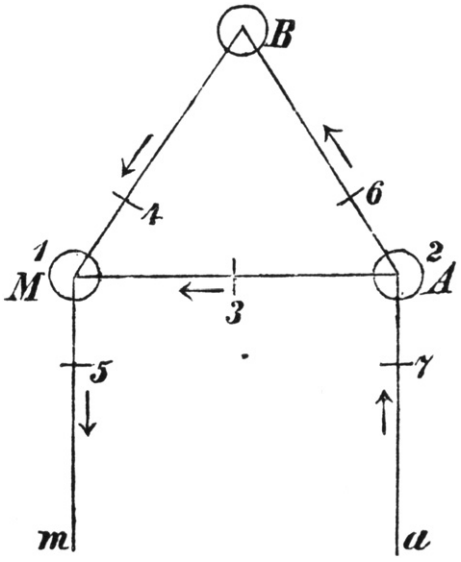

Figure 1.2.: Lichtheim's house diagram. A: word representation center, B: conceptual center, M: motor center. (Figure from Eling, 2011).

So, what progress has been achieved since the introduction of modern brain mapping techniques? Has neuroimaging supported the classical model?

First of all, the notion of Broca's and Wernicke's area as the *only* regions involved

in language processing and as regions involved *only in* language processing was no longer tenable. Broca's and Wernicke's area could no longer be considered the classical language centers where language processing exclusively takes place. An early PET study has clearly shown regions outside the classical language centers to be involved in language processing (Binder et al., 1997). Furthermore, on the eve of neuroimaging, it was accepted, in line with the classical model, that the left planum temporale (PT; the anatomically better-defined equivalent for Wernicke's area for scientists at the time), is a speech specific region (Galaburda, Sanides, & Geschwind, 1978; Geschwind & Levitsky, 1968). That is to say, the left planum temporale is driven by speech and speech only. However, an early auditory fMRI study by Binder, Frost, Hammeke, Rao, and Cox (1996) clearly refuted the idea that the left planum temporale is speech specific. These authors found it to be driven by acoustic as well as linguistic stimuli. Additionally, over time evidence accumulated showing that the left and the right PT both play a crucial role in sound and speech processing (e.g. Gandour et al., 2003; Jäncke, Wüstenberg, Scheich, & Heinze, 2002; Meyer, Alter, Friederici, Lohmann, & von Cramon, 2002; Meyer, Steinhauer, Alter, Friederici, & von Cramon, 2004). A recent meta-analysis by Vigneau et al. (2011) also confirmed this pronounced right hemisphere involvement in a variety of language-related processes. Hence, ignoring the right hemisphere in models of language and speech processing seems problematic at least.

As revolutionary the classical model has been at its introduction, from a present point of view neuroimaging has revealed a number of issues. Let us consider here one neuroanatomical and one linguistic problem. First, the model works with anatomical distinctions and predictions that time has proven to be problematic (for an elaborated analysis of this and other issues the reader is referred to Poeppel & Hickok, 2004). Neuroanatomical progress clearly showed that Broca's area and Wernicke's area are by no means homogeneously organized areas (e.g. Amunts & Zilles, 2012; Wise et al., 2001). Furthermore, as mentioned above, neuroimaging studies have pointed to regions outside the classical language areas, not restricted to the left hemisphere, that are involved in language and speech processing.

A further critical issue is that the classical model is linguistically substantially underspecified (Poeppel & Hickok, 2004). Language was considered to consist of merely two sub-components: production and comprehension. From a linguistic point of view this dramatically neglects the complexity of language and different concepts linguists use (e.g. phonology, syntax, semantics). This was recognized in the neuroimaging literature and resulted in a wide range of studies that aimed to study phonological, syntactic and semantic processing in isolation from one another (e.g. Friederici, 2002; Grodzinsky, 2000; Hagoort, 2005; Pulvermüller, 2010). Vivid "battles for Broca's region" have been fought (Grodzinsky & Santi, 2008) but a clear consensus is still lacking. In essence, the problem with the classical model – using concepts that are too coarse to represent the linguistic reality – is only shifted to another level in the aforementioned neuroimaging studies. Importantly, concepts like syntax or semantics "(...) are not monolithic,

and have many sublevels of representations (...)" (Boeckx, 2009, p. 159). Or as Van Lancker Sidtis (2006) wrote:

> "The proposed linguistic levels—phonetics, phonology, morphology, lexicon, syntax, and semantics—may be more useful as an educational and analytic heuristic than for describing the biology of human language and language disturbance. Their status as autonomous, explanatory elements in mental and cerebral processing is no longer clear. These terms may not reflect the language of the mind or the brain." (p. 279)

Taken together, Boeckx (2009) and Van Lancker Sidtis (2006) indicate that looking for brain modules that accomplish processes related to coarse linguistic concepts might not be a fruitful research program and establishing a direct link between linguistics and neuroscience is not as simple as a dominant fraction neuroscientific literature suggests. In more detail, Poeppel and Embick (2005) identify two major conceptual issues that impede describing a link between the two domains. First, the "Granularity Mismatch Problem" considers that linguists and neuroscientists operate with concepts of radically different granularity or coarseness. While linguists work with finely grained, well-defined concepts and explicitly formulated operations, typical neuroscientific concepts of language are by far broader and less explicitly defined. Second, the authors diagnose an "Ontological Incommensurability Problem": since the units of linguistic and neuroscientific computations have developed independently they are suggested to be incommensurable. Hence, a direct link between linguistic concepts (like syllable, clause, linearization or semantic composition) and neuroscientific concepts (like neuron, cortical column or oscillation) is not possible. To solve these problems Poeppel and Embick (2005) propose:

> "(...) spelling out the ontologies and processes in computational terms that are at the appropriate level of abstraction (i.e. can be performed by specific neuronal populations) such that explicit interdisciplinary linking hypotheses can be formulated. (...) [W]e recommend taking linguistic categories seriously and using them to investigate how the brain computes with such abstract categorical representations." (pp. 106 - 107)

For instance, the linearization of hierarchical structures is an important syntactic operation. It might also be involved in other linguistic or cognitive domains, like phonological or motor sequencing. The authors reason that studying linearization in different domains might therefore describe the problem at the appropriate level of granularity.

While recently published models tend to incorporate insight from recent neuroimaging studies, for instance with regards to the bihemispheric organization of language processing, and try to better define operational demands of speech processing (e.g. Hickok & Poeppel, 2007), it seems as if they have a hard time liberating themselves from old concepts. A framework adhering to the suggestions made by Poeppel and

Embick (2005) is a scheme by Ben Shalom and Poeppel (2008). The authors propose a taxonomy of language perception on two axes: first, along different linguistic aspects (phonology, syntax and semantics); second, along different linguistic operations (memorizing, analyzing and synthesizing). Those operations are present in all of the proposed aspects. Based on the linguistic operations, a coarse division of the cortex is proposed: the temporal lobe is linked to memorizing (storing and retrieving) linguistic items, the parietal lobe is associated with analyzing these items (accessing subparts of the items), the frontal lobe with synthesizing (combining and assembling) items. Within each operation, hence, within each mentioned lobe, an additional principle of organization is proposed. A superior-to-inferior gradient is assumed to exist, organizing the different aspects. In each lobe, phonetic/phonological information is localized most superiorly, followed by syntactic information in the middle and semantic representations most inferiorly. While the anatomical distinctions as well as the linguistic operations defined in this model are quite coarse, it definitely constitutes a step towards a testable model that might explain a variety of findings in a linguistically motivated manner, with a strong focus on computations required to process language and speech.

Another approach by Poeppel (2001, 2003), the "asymmetric sampling in time" (AST) hypothesis, focuses on more initial stages of speech perception in a bottom-up driven manner. Poeppel argues that low-level acoustic features of speech underlie lateralization of high-level speech-related processes in the brain. As noted, for instance, by Rosen (1992) information in the speech stream is transported over multiple temporal scales. The acoustic fine structure at a range of several milliseconds is represented by rapidly changing spectro-temporal patterns, for instance formant transitions. These rapidly changing patterns can constitute the perceptual differences between two phonemes. On the other hand, the acoustic envelope codes prosodic features of speech in the range of several hundred milliseconds, for instance intonation contour, tempo, rhythm, syllabic structure, and stress. Furthermore, as mentioned above and in contradiction to the classical model, neuroimaging studies have shown bihemispheric involvement of perisylvian cortex during speech perception. This involvement occurs lateralized depending on acoustical properties of the speech signal. This has been shown, for instance, by Belin et al. (1998) who reported that the perception of rapidly changing acoustic cues resulted in a left lateralized pattern of activity in the auditory cortex. Since rapid transitions are a critical part of the speech signal, the authors reason that this lateralization on perceptual level might underlie later, higher-level language-related brain functions.

Based on these and other observations, the AST framework was proposed (Poeppel, 2001, 2003). It argues that after an initial symmetrical representation of the speech signal in primary auditory cortex, different aspects of the incoming speech signal are preferentially processed in a lateralized manner in non-primary auditory cortex. While rapidly changing acoustic features (e.g. formant transitions) are preferentially processed left lateralized, slowly changing cues (e.g. prosodic features) are preferentially processed

right lateralized. A more detailed account of the AST framework can be found in the introduction of Study II (chapter 2.2.2).

Supporting evidence for the AST framework comes from fMRI studies using well-controlled, parametrically manipulated, acoustic, non-linguistic stimuli (e.g. Boemio, Fromm, Braun, & Poeppel, 2005; Schönwiesner, Rübsamen, & von Cramon, 2005) or brief speech stimuli (e.g. syllables; Britton, Blumstein, Myers, & Grindrod, 2009). The question arises whether longer, at least to some extent more speech-like, parametrically manipulated stimuli, for instance sentences, result in functional lateralization predicted by AST? Furthermore, does anatomical lateralization in the posterior temporal lobe predict perception when acoustic cues are manipulated?

Open question 1: Can the functional lateralization predicted by AST be shown with longer, more natural stimuli, in a novel parametric manipulation paradigm? Does anatomical lateralization in planum temporale predict perception in stimuli with altered acoustic information?

This question has been addressed in Study II.

1.3. Methods to study brain structure and function

We now turn to methodological issues, especially in the context of auditory and speech experiments, that have been addressed in this thesis.

1.3.1. Auditory functional magnetic resonance imaging

As already mentioned in chapter 1.1, functional magnetic resonance imaging is an indirect method to study brain functions. By utilizing neurovascular coupling, changes in neuronal activity are deduced from changes in the ratio of oxygenated and deoxygenated hemoglobin. This ratio can be observed in the blood-oxygen-level-dependent (BOLD) signal (Ogawa et al., 1990). To record this signal, volunteers are placed inside a scanner that produces a strong permanent magnetic field. Additional weak magnetic fields (gradients) and radiofrequency pulses are used to measure differential magnetic response of oxygenated and deoxygenated hemoglobin. This method produces three-dimensional functional maps of the brain with a relatively high spatial resolution (millimeter range). One severe drawback is the acoustic noise that is emitted as the gradient coils are switched that can reach levels up to 130 dB. Therefore, auditory stimulation during image acquisition is severely disturbed by acoustic scanner noise. However, the very slow and delayed trajectory of the hemodynamic response, otherwise a pivotal disadvantage of fMRI, makes it possible to present auditory stimuli in the absence of scanner noise. This delay is as follows: when auditory stimulation evokes firing of auditory cortex neurons, the hemodynamic response slowly increases and reaches its maximum at approximately

five seconds after the stimulation. This indicates that the temporal resolution of fMRI is far from optimal (in the range of several seconds). These temporal characteristics have been exploited by acquisition procedures which present auditory stimuli and perform a delayed image acquisition in pauses between stimuli. The "sparse temporal sampling" approach (Hall et al., 1999) samples one brain image per cycle and has been further advanced into the "clustered-sparse" imaging approach that samples multiple brain volumes per cycle (Schmidt et al., 2008; Zaehle, Schmidt, et al., 2007). These techniques enabled scientists to conduct reasonable auditory fMRI studies in the first place. For a more detailed introduction into auditory fMRI scanner protocols the reader is referred to the introduction of Study I (chapter 2.1.2).

Based on theoretical considerations, it has been assumed that long silent pauses between consecutive image acquisitions enable a better measurement of brain responses evoked by auditory stimuli. As these theoretical considerations heavily rely on multiple processes (e.g. several hemodynamic responses evoked by different stimuli in rapid succession, potential saturation effects) interacting in a complex manner, the actually resulting modulation of the hemodynamic response is difficult to predict. Interestingly, the question of the optimal timing in modern auditory fMRI protocols has not yet been tested empirically.

Open question 2: Can the theoretically assumed optimal timing for auditory fMRI protocols be supported by empirical evidence?

This question has been addressed in Study I.

1.3.2. Structural brain imaging

Beside brain function, magnetic resonance imaging can depict the brain's structure, as well. Since different tissue types (gray matter, white matter, cerebrospinal fluid) exhibit differential magnetic properties, they also show different levels of intensity in brain images acquired with structural magnetic resonance imaging. Images can be analyzed to characterize cortical and subcortical structures and, subsequently, to correlate measurements of structure with external variables. This can be achieved via several analysis approaches. Most prominently, voxel-based morphometry (VBM; Ashburner & Friston, 2000; Wright et al., 1995) has produced a diversity of structural insight over the last decade. In recent years, more elaborated surface-based morphometry (SBM) approaches have become increasingly accessible and popular (Dale, Fischl, & Sereno, 1999; Fischl, Sereno, & Dale, 1999). In this thesis, SBM was used in the studies II and III.

Voxel-based morphometry

Before the introduction of automated morphometry algorithms, the only way to study neuroanatomy in-vivo was to manually label brain structures of interest on anatomical brain images. This is a time-consuming and error-prone endeavor, which requires considerable neuroanatomical knowledge and multiple raters to ensure objectivity. Therefore, automatic approaches constituted a possibility to run objective anatomical studies on a large sample in a reasonable amount of time with the possibility to analyze the whole brain, not only a few structures. One common approach is the VBM procedure by Wright et al. (1995), which is implemented in the SPM software package (Ashburner & Friston, 2000).

In a first step, SPM's VBM procedure identifies gray matter in the structural brain volume of each subject (segmentation). These gray matter maps are registered onto a common standard brain and then are spatially smoothed. Per subject, this results in a three-dimensional map. In this map, each voxel (hence the name) represents a structural measurement. In the classical VBM framework these measures were referred to as "gray matter density" or "gray matter probability". These measurements are relatively difficult to interpret as they lack a clear neuroanatomical equivalent. Recent improvements take the amount of scaling at the registration step into account and correct for it. The resulting measurement represents gray matter volume or "tissue volume per unit volume of spatially normalized images" (Ashburner, 2009, p. 1166). Finally, these values are entered into a voxelwise mass-univariate statistic in order to identify group differences, correlations with external variables or changes over time within a person. VBM poses an objective procedure with relatively low computational demands that is able to process the whole brain, i.e. cortical and subcortical structures, in one analysis. However, several critical disadvantages exist. Results are is highly sensitive to misclassification in the segmentation step and to misalignment in the registration step (Ashburner, 2009). To mitigate these shortcomings, strategies to enhance segmentation and registration have been developed (Ashburner & Friston, 2005). Even more crucial, VBM's gray matter volume might be confounded by other measurements. Differences in gray matter volume can occur as a result of differences in folding pattern (resulting in a different cortical surface area), differences in cortical thickness or due to processing artifacts (Ashburner, 2009). As cortical volume is determined by the product of cortical thickness and surface area, and VBM is not able to measure those indicators separately, other approaches which can achieve this, have been devised, for instance surface-based morphometry.

Surface-based morphometry

A different reasoning is behind surface-based approaches implemented, for instance, in Freesurfer (Dale et al., 1999; Fischl, Sereno, & Dale, 1999). These procedures consider the inherent two-dimensional topology of the cerebral cortex. The cortex can be regarded as a folded, two-dimensional sheet. For example, let us consider two

neighboring but functionally entirely distinct gyri; in a 3-D volumetric representation these gyri might touch each other, however in a 2-D surface representation they might be located several centimeters apart. Although surface-based representations can improve visualization, e.g. in sulci, this is not merely a visualization issue. In the case of volumetric smoothing (as performed in VBM) data from locations that are far apart on the cortical sheet can be mixed together, which decreases data quality. Taken together, surface-based representations are more appropriate to represent the complexity of cortical architecture in the context of in-vivo imaging studies (Van Essen, Drury, Joshi, & Miller, 1998).

Hence, Freesurfer's SBM aims to reconstruct models of the white matter surface (the interface between white and gray matter) and the pial surface (the border between gray matter and cerebrospinal fluid) in each subject's native space (cf. Figure 1.3). This way, the highly individual cortical folding pattern of each subject can be captured.

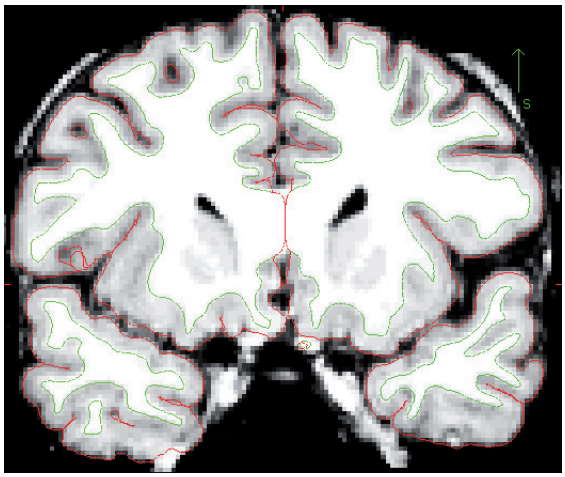

Figure 1.3.: Volumetric T1-weigted brain image with automatically generated surface models. Green: white surface; red: pial surface.

Cortical thickness can be calculated from these surfaces (for an example of a projection of cortical thickness of the entire cortex see Figure 1.4). For each point on the surface, cortical thickness is defined as the shortest distance between the white and the pial surface models. Furthermore, the reconstructed folding patterns can be parcellated into gyral and sulcal regions of interest (ROIs; cf. Figure 1.5).

This enables the researcher to compare measurements like cortical thickness, surface area or volume within a ROI across a group of subjects. Alternatively surfaces can be registered onto a standard template surface to perform vertexwise comparisons of these cortical measurements. In any case, metrics resulting from this approach are fairly straightforward to interpret, as they clearly represent a geometrical property of the cortex and are stated in standard units (e.g. cortical thickness in mm, surface area in mm^2). In a separate analysis, subcortical structures can be measured with Freesurfer as well. It has to be noted that SBM is far from perfect. The quality of reconstructed surfaces

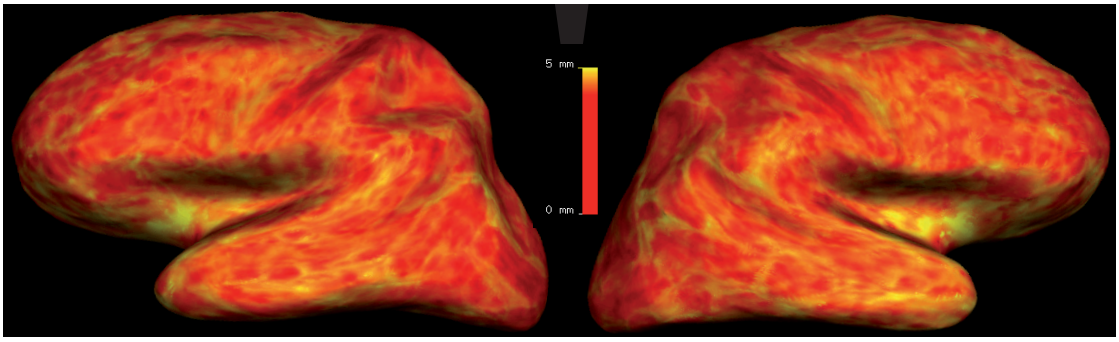

Figure 1.4.: Cortical thickness of a single person over the entire cortex projected onto an inflated surface representation.

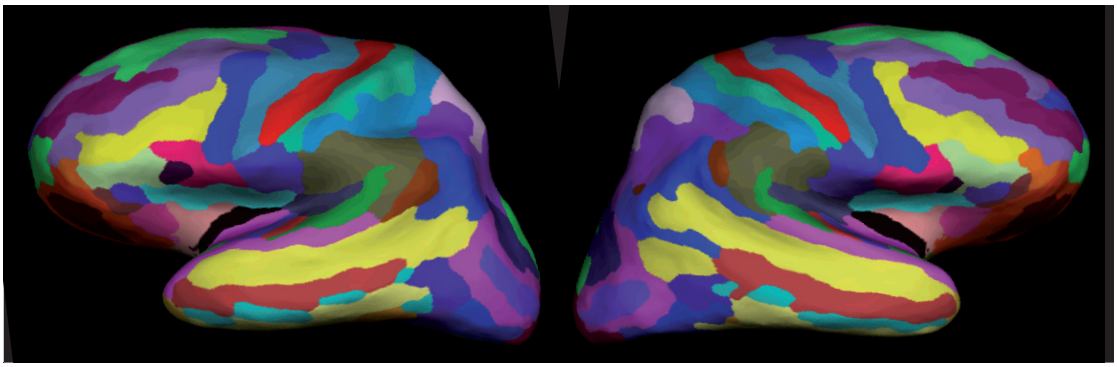

Figure 1.5.: Automatic parcellation of the cortex of a single person into regions of interest.

relies on the quality of the entered images. In special cases, a manual interference with the algorithm can be necessary.

Voxel-based morphometry vs. Surface-based morphometry

From a conceptual point of view SBM outperforms VBM on several levels. SBM constructs surface models that follow the individual folding pattern of a subject, whereas VBM is blind to any idiosyncrasies in cortical architecture. Furthermore, in Freesurfer's SBM spatial accuracy is finer than the voxel resolution, therefore, group differences far below the voxel size (i.e. smaller than one millimeter) can be analyzed. On the other hand, VBM in SPM is characterized by a relatively fast computation (while processing one subject with Freesurfer's SBM takes around 24 hours, it takes around 15 minutes in VBM; this makes VBM around 38 times faster) and an easy to use graphic user interface. Especially in studies with large samples, computational time might be a point that needs to be considered. The derived measurement in VBM can be neurobiologically difficult to interpret and a confound of several other measurements (Panizzon et al., 2009; Voets et al., 2008; Winkler et al., 2010). Via Freesurfer's SBM a wider array of measurements can be calculated separately, i.e. cortical thickness, surface area, volume

and more sophisticated measures like the gyrification index. Furthermore, SBM to a lesser extent relies on the registration of single subjects to a standard template, a critical and error-prone step in VBM (Mietchen & Gaser, 2009). Several empirical studies indicate that SBM approaches depict neuroanatomy more sensitively (Li et al., 2012; Pereira et al., 2011) and reliably (Clarkson et al., 2011; Chalavi, Simmons, Dijkstra, Barker, & Reinders, 2012) with a finer spatial resolution than VBM (Cerasa et al., 2012). Nevertheless, in some instances, e.g. in learning and plasticity studies, VBM might be more sensitive to subtle changes in brain anatomy over time, albeit decreased spatial resolution to localize these changes (Mietchen & Gaser, 2009).

It has to be noted that the MR images fed into the VBM and the SBM processing tools, are identical. Increase in the sensitivity of both procedures will be achieved by improving structural imaging. Images with a finer spatial resolution and higher image quality will result in more accurate measurements of brain anatomy(May & Gaser, 2006).

Taken together, while VBM studies have lead to an impressive amount of insight into neuroanatomical correlates of behavior, in many contexts, the use of recently developed SBM methods is preferable over the use of VBM methods to characterize brain structure.

While MR-based measurements of brain structure and function dominated the last decades, an electrophysiological technique, the electroencephalogram (EEG), was introduced long before that, over 80 years ago. EEG made it possible to noninvasively study brain function in humans,

1.3.3. Auditory evoked potentials

The electroencephalogram is able to directly measure electrical brain activity via electrodes placed on the scalp of a volunteer. The first attempts to record electrical currents emitted by the brain have been made by the English physiologist Richard Caton in 1875. He measured those currents on the brain surface of animals. The human EEG was introduced by the German psychiatrist Hans Berger in 1929. He was the first to describe differences between the EEG at sleep and awake (Bear, Connors, & Paradiso, 2007). The EEG gained popularity, as it was, until the introduction of PET and fMRI, the only method in cognitive neuroscience to noninvasively study brain activity. Classically, it was not considered a brain mapping technique, since, due to the inverse problem, it was not possible to produce three-dimensional maps of brain function. Therefore, this method was overshadowed by the alluring brain maps produced with PET and fMRI. However, mathematical advances that are able to estimate the source of electrical activity in the brain from recordings of activity on the scalp level dramatically increased the potential that EEG has for neuroscientific studies. Moreover, EEG's excellent temporal resolution, in comparison to PET or fMRI, enables this method to study brain processes at the appropriate temporal resolution of milliseconds (Jäncke, 2005). The fact that EEG recordings no not emit acoustic noise makes this device ideal for auditory stud-

ies. To compensate for the coarse spatial resolution, fruitful attempts to simultaneously record fMRI and EEG have been made, resulting in data that represents brain function with an outstanding spatial as well as temporal resolution (e.g. Bonmassar, Anami, Ives, & Belliveau, 1999; Mulert et al., 2005; Scarff et al., 2004). For an review of the history of simultaneous EEG-fMRI recordings see Herrmann and Debener (2008).

Electrical currents can be measured on the scalp because they spread from neurons through the cortex, the meninges and the skull. On the neuronal level, synchronized firing of parallel organized cortical neurons produces extracellular field potentials. If the mass of synchronized neurons is large enough, the fields propagate to the scalp and can be measured with electrodes (Brandeis, Michel, & Amzica, 2009). Auditory evoked potentials (AEPs) can be measured by repeatedly presenting an auditory stimulus to a volunteer and then averaging the time locked signal. With this procedure, electrical brain activity that is evoked by the stimulus can be separated from other, unrelated signals (since AEPs are a standard procedure in neuroscience, used for several decades, the reader is referred to chapter 19 in Bear et al. (2007) or chapter 8 in Jäncke (2005) for a more detailed account).

The auditory N1 component is an AEP that is related to basic auditory processing (Meyer, Baumann, & Jancke, 2006; Näätänen & Picton, 1987). It is influenced by basic stimulus properties, for instance loudness (Mulert et al., 2005). Furthermore, the N1 changes as a function of auditory experience, for instance professional musical training (Baumann, Meyer, & Jäncke, 2008; Pantev, Roberts, Schulz, Engelien, & Ross, 2001). Based on intracortical recording studies as well as source estimation experiments, it is well established that the N1 is generated in sources along the supratemporal plane (e.g. Baumann et al., 2008; Godey, Schwartz, de Graaf, Chauvel, & Liégeois-Chauvel, 2001; Liégeois-Chauvel, Musolino, Badier, Marquis, & Chauvel, 1994; Yvert, Crouzeix, Bertrand, Seither-Preisler, & Pantev, 2001; Zaehle, Jancke, & Meyer, 2007).

However, little is known about the relationship between macroscopic neuroanatomy and the N1 amplitude. It has been shown that the N1 amplitude is stable over time within a person (Roth, Kopell, Tinklenberg, Huntsberger, & Kraemer, 1975; Walhovd & Fjell, 2002). Nevertheless, it can depict plastic brain changes of brain function. For instance, it is modulated by professional musical training (Baumann et al., 2008; Pantev et al., 2001). Since this parallels the macroscopic neuroanatomical characteristics of stability and plasticity, a relationship between macroscopic neuroanatomy, i.e. cortical thickness and surface area, and the N1 amplitude is suggested. For instance, brains with a thicker cortex in the supratemporal plane might show enhanced N1 amplitudes (for a more detailed elaboration on this topic the reader is referred to the introduction of Study III (chapter 2.3.2)).

The question arises whether measurements of cortical anatomy derived from surface-based morphometry approaches (e.g. cortical thickness, surface area) are able to predict AEPs.

Open question 3: Are the characteristics of macroanatomical cortical features in the supratemporal plane a predictor of the auditory N1 amplitude?

This question has been addressed in Study III.

1.4. Aims and significance

To conclude the introduction, the main aims of the three empirical studies, as well as their significance and novelty, will be summarized. Each of the studies was designed to address one of the open questions raised in the introduction:

- **Open question 1:** Can the functional lateralization predicted by AST be shown with longer, more natural stimuli, in a novel parametric manipulation paradigm? Does anatomical lateralization in planum temporale predict perception in stimuli with altered acoustic information?

  Study II (chapter 2.2, Liem, Hurschler, Jäncke, & Meyer, 2013) aimed to asses structural and functional asymmetry in auditory-related cortex dependent on the availability of acoustic cues in speech. By using an innovative stimulus manipulation approach, this study struck a balance between the requirements to use ecological valid, as well as properly controlled stimulus material. As predicted, the results yielded acoustic-dependent functional lateralization in the planum temporale and posterior superior temporal gyrus, but not in primary auditory cortex. Furthermore, lateralization of planum temporale thickness predicted perceptual performance.

- **Open question 2:** Can the theoretically assumed optimal timing for auditory fMRI protocols be supported by empirical evidence?

  Study I (chapter 2.1, Liem, Lutz, Luechinger, Jäncke, & Meyer, 2012) aimed to optimize the timing in auditory fMRI settings. This study was the first to empirically examine timing issues in this context and resulted in a remarkable improvement of the "clustered-sparse" imaging protocol. The results from this study enable researcher to design auditory fMRI studies with increased statistical power, which makes it possible to perform more conclusive experiments.

- **Open question 3:** Are the characteristics of macroanatomical cortical features in the supratemporal plane a predictor of the auditory N1 amplitude?

  Study III (chapter 2.3, Liem, Zaehle, Burkhard, Jäncke, & Meyer, 2012) aimed to examined the neuroanatomical basis of the auditory electrophysiological signal. It is the first study to report a correlation between cortical thickness in the supratemporal plane and auditory evoked potentials.

2. Empirical Part

2.1. Reducing the Interval Between Volume Acquisitions Improves "Sparse" Scanning Protocols in Event-related Auditory fMRI.

Franziskus Liem, Kai Lutz, Roger Luechinger, Lutz Jäncke, Martin Meyer

Published in Brain Topography, 25(2), 182-193. doi: 10.1007/s10548-011-0206-x

2.1.1. Abstract

Sparse and clustered-sparse temporal sampling fMRI protocols have been devised to reduce the influence of auditory scanner noise in the context of auditory fMRI studies. Here, we report an improvement of the previously established clustered-sparse acquisition scheme. The standard procedure currently used by many researchers in the field is a scanning protocol that includes relatively long silent pauses between image acquisitions (and therefore, a relatively long repetition time or cluster-onset asynchrony); it is during these pauses that stimuli are presented. This approach makes it unlikely that stimulus-induced BOLD response is obscured by scanner-noise-induced BOLD response. It also allows the BOLD response to drop near baseline; thus, avoiding saturation of BOLD signal and theoretically increasing effect size. A possible drawback of this approach is the limited number of stimulus presentations and image acquisitions that are possible in a given period of time, which could result in an inaccurate estimation of effect size (higher standard error). Since this line of reasoning has not yet been empirically tested, we decided to vary the cluster-onset asynchrony (7.5, 10, 12.5, and 15 s) in the context of a clustered-sparse protocol. In this study sixteen healthy participants listened to spoken sentences. We performed whole-brain fMRI group statistics and region of interest analysis with anatomically defined regions of interest (auditory core and association areas). We discovered that the protocol, which included a short cluster-onset asynchrony (7.5 s), yielded more advantageous results than the other protocols, which involved longer cluster-onset asynchrony. The short cluster-onset asynchrony protocol exhibited a larger number of activated voxels and larger mean effect sizes with lower standard errors. Our findings suggest that, contrary to prior experience, a short cluster-onset asynchrony is advantageous because more stimuli can be delivered within any given period of time. Alternatively, a given number of stimuli can be presented in less time, and this broadens the spectrum of possible fMRI applications.

2.1.2. Introduction

Despite the delayed temporal characteristics of the hemodynamic response, functional magnetic resonance imaging (fMRI) is a major tool in present cognitive neuroscience. As the scanner produces auditory noise, a variety of problems emerge, especially in the context of auditory fMRI studies (Amaro et al., 2002; Moelker & Pattynama, 2003). Major disadvantages include: acoustical overlap between scanner noise and stimulus, and an enhancement of the apparently "silent" baseline condition in continuous protocols. The overlap between scanner noise and stimulus presentation makes it difficult to perceive auditory stimuli in their full complexity, especially if subtle stimulus manipulation is applied. Perceived spectral characteristics of stimuli are altered by ambient scanner noise that can be as loud as 130 dB. Furthermore, a research participant in the scanner may have to be more attentive, in order to perceptually separate the auditory stimulus from background scanner noise. On a physiological level, auditory scanner noise can lead to saturation of neuron populations in auditory fields because this intense perpetual noise excessively drives auditory cortex activity. Scanner noise induces a BOLD response in auditory-related cortex areas during trials without a proper auditory stimulus. Most interestingly, this appears to happen differentially for the left and the right hemisphere (Herrmann, Oertel, Wang, Maess, & Friederici, 2000; Tamer, Luh, & Talavage, 2009; Schmidt et al., 2008) and in a nonlinear manner (Talavage & Edmister, 2004). The degree of nonlinearity varies between left and right hemisphere (Hu et al., 2010). The additional auditory input leads to an "inflated" baseline condition, which reduces the possible range of stimulus-induced BOLD response (more detailed accounts of these and further problems have been given for example by Eden, Joseph, Brown, Brown, & Zeffiro, 1999; Gaab, Gabrieli, & Glover, 2007a, 2007b; Hall et al., 1999).

To overcome these constraints, several groups have published groundbreaking techniques, which are standard in today's auditory fMRI (Eden et al., 1999; Edmister, Talavage, Ledden, & Weisskoff, 1999; Hall et al., 1999; Talavage, Edmister, Ledden, & Weisskoff, 1999). A variety of names exists to date for similar approaches: Edmister et al. (1999) and Talavage et al. (1999) presented "clustered volume acquisition"; Hall et al. (1999) called their approach "sparse temporal sampling"; Eden et al. (1999) published the "behavior interleaved gradients technique". In the present paper we use the term "sparse design" for aforementioned approaches and "clustered-sparse design" for an extension of this design (Schmidt et al., 2008; Zaehle, Schmidt, et al., 2007). "Clustered volume acquisition" refers to the temporal clustering of several slices within one volume and is not to be confused with the clustered-sparse protocol, which refers to the clustering of volumes within one trial.

The "sparse" temporal acquisition scheme reduces the scanner's acoustical noise influence by acquiring only one functional image per trial. When repetition time (TR) is long (around 10 – 14 s), the preceding trial's scanner-evoked BOLD signal returns

close to baseline. In this scheme, auditory stimuli are delivered during the silent pause between two image acquisitions. This allows for unobstructed stimulus perception. In addition to these benefits, sparse acquisition schemes show a higher SNR because T1 magnetization can fully recover prior to each image acquisition. This is impossible in continuous protocols. Notwithstanding these undisputable advantages this approach has some drawbacks. The total duration of an experiment is increased and image acquisition needs to be timed around the peak of the BOLD response. Timing is not a pressing issue in block designs where the stimulus-evoked BOLD response almost reaches a steady state at the plateau (e.g. when brief tones or syllables are repeatedly presented for approximately ten seconds followed by an acquisition of one functional volume). However, whenever single-stimulus presentation in the context of an event-related design is desired, timing is a key issue (cf. Figure 2.1). For example, timing is a key issue, when investigating the perception of slow modulations in auditory stimuli with durations in the range of a few seconds, such as, prosody in sentences, or melody in brief excerpts of music.

To ensure that the BOLD response's peak is sampled, the clustered-sparse temporal acquisition (CTA) protocol has been devised (Schmidt et al., 2008; Zaehle, Schmidt, et al., 2007). Derived from the sparse protocol, the clustered-sparse scheme allows for the collection of a cluster of (usually three) consecutive functional volumes per trial. As a result, the likelihood of recording the hemodynamic response's peak increases, as does the number of acquired images and therefore the number of observations. This makes the clustered-sparse protocol superior to the sparse approach with respect to statistical power. Especially in single subject analyses, this superiority becomes manifest in more precisely estimated effects, namely, beta-values (Zaehle, Schmidt, et al., 2007). The duration of the silent pause in CTA designs is determined by the cluster-onset asynchrony (COA, the time between two consecutive cluster-onsets).

Interestingly, there does exist an approach similar to the clustered-sparse scheme: silent gradient protocols (Mueller et al., 2011; Schmithorst & Holland, 2004; Schwarzbauer, Davis, Rodd, & Johnsrude, 2006) also sample data by using a cluster of several volumes. Clusters are separated by silent intervals during stimulus presentation. In contrast to our approach, longitudinal magnetization is held constant via silent slice-selective excitation pulses. As a result, the T1-decay-related signal-to-noise ratio (SNR) improvement (for at least the first image per cluster), which the clustered-sparse protocol benefits from, is absent.

Especially in studies focusing on auditory-related areas silent protocols, in comparison to conventional continuous acquisition protocols, have been shown to be beneficial in terms of SNR and effective power (Gaab et al., 2007a, 2007b; Hall et al., 1999; Schmidt et al., 2008). However, a slow timing and the resulting inflated duration of the scanning session can make such studies a tedious and gruelling experience for participants. Therefore, the aim of this study is to empirically improve the timing setup by varying the COA. Compared with long COAs (e.g. 15 s), short COAs (e.g. 7.5 s) lead to a notable

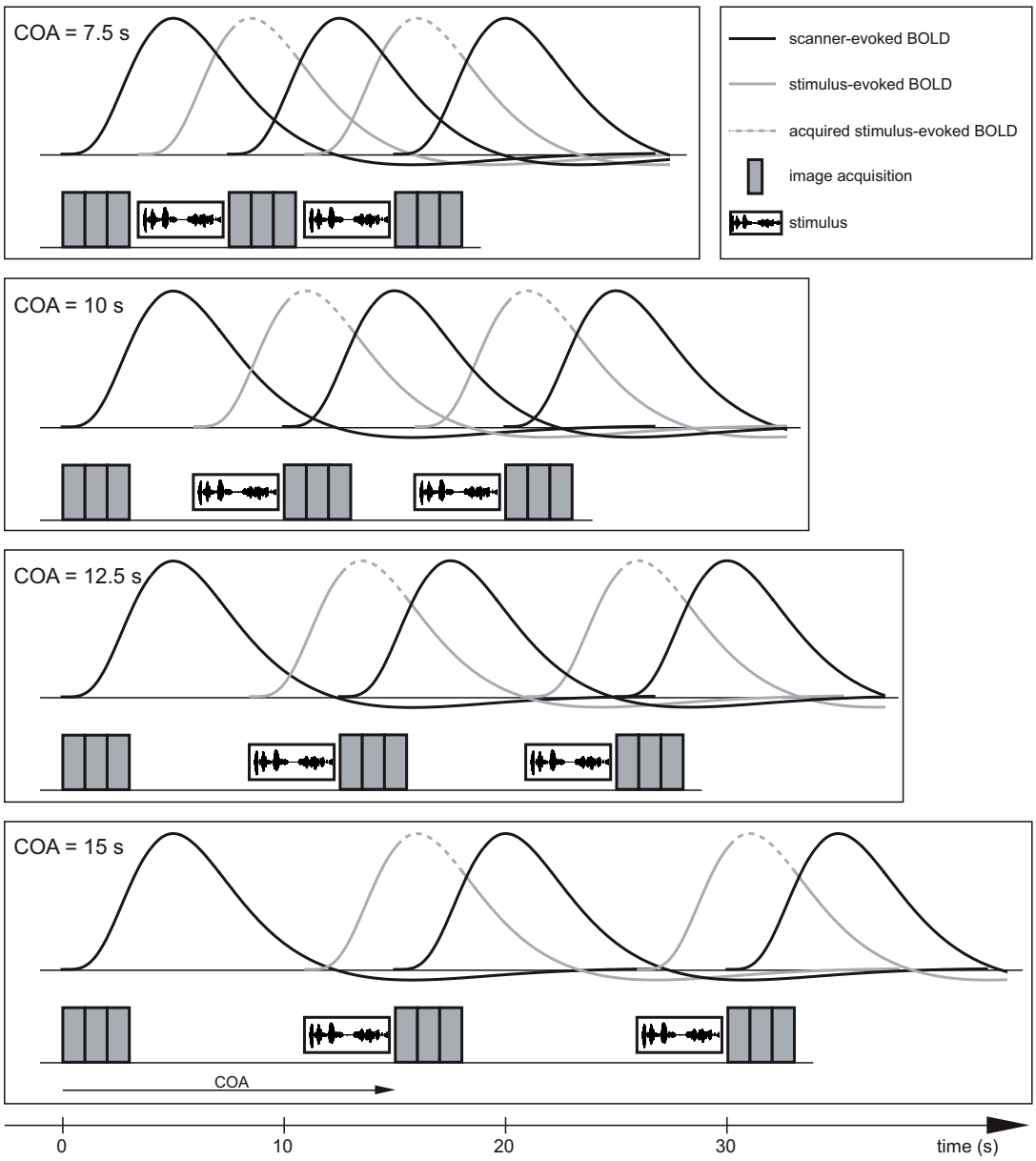

Figure 2.1.: Assumed BOLD response evoked by auditory stimuli and auditory scanner noise, respectively.

increase in the amount of acquired data in a given period of time. This should increase the accuracy of the parameter estimates, as manifested in the parameter estimate's lower standard error of the mean (SEM).

The SEM is of major interest because it directly influences t-values and, therefore, the statistical significance. Generally spoken, if two designs show identical beta-values but different standard errors the t-value is higher in the design with the smaller standard error (Mechelli, Price, Henson, & Friston, 2003). We applied an approach introduced by Mechelli et al. (2003) to segregate the SEM ($(\sigma\ ^2 c^T (X^T X)^{-1} c)^{1/2}$) into error variance

(σ^2) and design variance ($c^T(X^TX)^{-1}c$). The design variance is solely depending on the design matrix and the contrast and represents a measurement of the variance of the explanatory variables and their correlations. The error variance is variance in the data that cannot be explained by the model. While short COAs allow for the presentation of a large number of trials, long COAs permit a clear separation of stimulus-evoked and scanner-noise-evoked BOLD response, as well as a return of BOLD response to baseline level. In addition, saturation effects are not expected to occur and, thus, do not influence the signal. This should result in a wider dynamic range of the BOLD signal (and a larger potential influence of auditory stimuli on the BOLD signal) and, therefore, in higher effect sizes (beta-values). Whereas the design variance is expected to be favourably influenced (reduced) by the higher number of images in a short COA setting, this effect should vanish if the amount of analyzed data is kept equal across the conditions by analyzing only part of the samples in short COA settings. In the present study, this was done by analyzing only the first 30 trials of each condition. In contrast, effect sizes should be fairly unaffected by such an analysis. Error variance is also not expected to be influenced by the number of trials. However, neurophysiological processes, for instance saturation effects or changes in the shape or temporal characteristics of the hemodynamic response, might render the model less appropriate, and therefore increase error variance. Whether decreases in design variance translate to decreases in SEM depends on error variance.

The current investigation focuses on event-related clustered-sparse designs, which shall exam hemodynamic response to processing of stimuli spanning over several seconds, namely spoken sentences and music stimuli, because these designs can elucidate neuronal mechanisms supporting slow acoustic modulation.

For this purpose, four differential COA settings were implemented (cf. Figure 2.1): 7.5, 10, 12.5 and 15 s (the latter was tested in prior clustered-sparse fMRI studies (Schmidt et al., 2008; Zaehle, Schmidt, et al., 2007)). To evaluate to what extent and how differential COA settings may influence the ability to identify different brain responses to stimuli that varied in loudness, we presented participants with sentences of two different intensities. Prior studies have convincingly shown an increase in the number of significantly activated voxels or percent signal change or both in auditory-related cortex areas as a result of increasing stimulus intensity (e.g. Brechmann, Baumgart, & Scheich, 2002; Hart, Hall, & Palmer, 2003; Jäncke, Shah, Posse, Grosse-Ryuken, & Müller-Gärtner, 1998; Mulert et al., 2005). In the present study, the variations in intensity were implemented merely as a vehicle to show differences in the settings' sensitivities.

It is assumed that the preceding trial's scanner-evoked BOLD response has less influence on the current image at long COAs because the signal has time to return to, or near baseline (cf. Figure 2.1; Hall et al., 2000). Therefore, we expect a systematic increase in effect size (beta-values) at longer COA settings. Notably, a short COA setting provides more data within a given period of time, which reduces design variance. If all COAs yield equal error variance, SEM will decrease as a function of decreasing design

variance, which leads to more precisely estimated effect sizes. Our aim is to find a COA setting that balances adverse impact on effect sizes and their estimation accuracy.

2.1.3. Materials and methods

Participants

Sixteen subjects (eight female) took part in this experiment. Participants were between 20 and 27 years old (M = 23, SD = 2). They were screened for hearing impairments, tinnitus, dyslexia, neurological and neuropsychological history. Participants were also asked if they had any metal implants or devices in their bodies. Subjects had normal or corrected vision. All subjects were German or Swiss German native speakers and right-handed according to the Annett questionnaire (Annett, 1992). They gave written informed consent and were paid for their participation. This study was approved by Canton Zurich's Ethics Committee (application E-40/2009).

Stimuli, experimental conditions and task

In this study we presented spoken German sentences to participants whilst they were placed in an MR scanner. The sentences were spoken by a trained female speaker and were recorded in a soundproof chamber at the University of Zurich Phonetics Lab. Each of the 146 sentences lasted on average about 3.0 s (SD = 0.4 s). Each sentence was presented only once during the course of the experiment (example below).

Sentences were delivered across four randomized runs (= four different conditions), each of which varied in COA (7.5, 10, 12.5 and 15 s). The COA is the single within-subjects factor of interest in this experiment. One run lasted 7.5 min and was composed of the maximum number of trials possible, depending on the COA (cf. Table 2.1).

Two other within-subjects factors of no direct interest were varied in this experiment: sentence intensity and sentence accent. In order to measure the COA setting's influence on the ability to detect differences between two classes of stimuli, sentences were presented pseudorandomly with either of two sound pressure levels (SPLs). By using the Praat software (v5.1.09; http://www.fon.hum.uva.nl/praat/) we set the sentences' mean intensity to 60 and 80 dB SPL, respectively. The stimuli were recorded with accents that were placed either on the first, or the second part of the sentence (for example: "Laura empfiehlt Martin, den Computer zu kaufen." / "Laura advises Martin, to buy the computer." The underline indicates the possible position of sentence accent). To control for participants' attention, they were asked to indicate by button press whether the sentence they just heard had a sentence accent on the first or the second noun phrase. Manipulation of intensity and emphasis was conducted in an orthogonal, randomized manner (within one run: 50% of sentences were presented at 60 dB, 50% at 80 dB; of each intensity group 50% with emphasis on the first part, 50% with emphasis on the second part). As a baseline measurement, empty trials were also included at a lower rate (cf. Table 2.1).

Data acquisition

An event-related clustered-sparse fMRI design was employed in this study (Schmidt et al., 2008; Zaehle, Schmidt, et al., 2007). Via MR-compatible headphones with an incorporated piezoelectric auditory stimulation system, one auditory stimulus per trial was binaurally presented in an interval devoid of auditory scanner noise. Throughout the experiment, sentence onset was four seconds prior to acquisition onset (cf. Figure 2.2). A fixation cross preceded the presentation of each sentence. The fixation cross was projected onto a screen and could be seen through a mirror mounted on the head coil. Subsequently, three functional volumes were recorded, each with an acquisition time of 1000 ms. During this interval, subjects indicated the noun phrase on which the accent was present via a button press with either their right index finger, or with their right middle finger (index finger for accents on the first noun phrase, middle finger for second noun phrase). During empty trials participants were asked to randomly press a button once, which enabled us to control for motor activity. The headphones' volume was calibrated with an SPL-meter prior to each session.

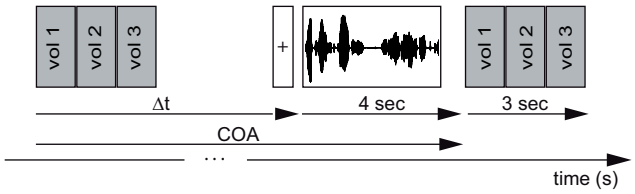

Figure 2.2.: Sequence of one trial. Gray squares: image acquisition. The interval between acquisition onset and the following stimulus onset varies with COA: Δt = COA - 4 s.

Data was collected on a Philips 3T Achieva whole-body MR unit (Philips HealthCare, Best, The Netherlands) that is equipped with an eight-channel Philips head coil. Functional time series were collected from 16 transverse slices covering the entire perisylvian cortex with a spatial resolution of 2.75 x 2.75 x 2.75 mm^3, using a single-shot, gradient-echo planar sequence (EPI acquisition matrix 80 x 80 voxels, field of view (FOV) = 220 mm, echo time (TE) = 35 ms, flip angle (FA) = 68°, SENSE factor = 2). Volume acquisition time of each EPI scan was 1000 ms. Cluster-onset asynchrony was systematically varied across the four runs (7.5, 10, 12.5 and 15 s, see Stimuli, experimental conditions and task section) but was kept constant within each run. Additionally, one whole-brain EPI (60 slices) was recorded prior to the experiment, so as to improve the spatial normalisation process according to an established procedure.

Furthermore, a standard 3D T1-weighted scan with 1 x 1 x 1 mm^3 spatial resolution (160 sagittal slices, FOV = 240 mm, TE = 3.7 ms, TR = 8.1 ms, FA = 8°) was collected, in order to obtain individual anatomical regions of interests (ROIs).

Data analysis

Behavioural data analysis and ROI statistics were performed using PASW Statistics 18.0 (SPSS Inc.).

Whole-brain analysis [1]

Since a whole-brain group analysis is standard in fMRI experiments, we included this analysis; however, the origin of a significant effect is more systematically to evaluate in a post-hoc ROI analyses of effect sizes and its SEM. Analysis of fMRI data was carried out using SPM8 (Wellcome Department of Cognitive Neurology, London). To account for movement artefacts, the functional images were realigned to the first volume. Each run was entered as a separate session. Since the functional brain volumes where comprised of only 16 slices, the realigned images were co-registered with the whole brain EPI. This permitted an overall improved normalisation. The whole-brain EPI was normalised onto the SPM8 EPI template. Resulting spatial normalisation parameters were applied to all functional volumes. This transformed them into MNI space. Finally, the images were smoothed with an FWHM kernel of 5 x 5 x 5 mm^3 (Buchsbaum et al., 2005).

After pre-processing the data, a General Linear Model (GLM; subject level) was separately set up for each run (= COA setting). Sentence-events were entered as two separate conditions (60 dB and 80 dB). Due to the low number of sampling points, a boxcar function (first order, window length = 3 s) was modelled for each trial. In accordance with the approach established by Zaehle, Schmidt, et al. (2007), two regressors of no interest were included to account for the T1-decay along the three consecutive volumes. Three contrasts were calculated: an auditory default contrast (all auditory events vs. empty trials) and two direct comparisons: 80 dB vs. 60 dB and 60 dB vs. 80 dB. For each COA, individual contrast images were subjected to a random-effects second level analysis (one-sample t-test against zero for all three abovementioned first level contrasts). Family wise error (FWE) correction was applied to the resulting statistical parametric maps. For each COA, suprathreshold voxels at the 80 dB vs. 60 dB contrast in the temporal lobe were counted and averaged across the two hemispheres for better statistical power.

ROI analysis To elaborate on effect sizes (mean beta-values) and distinct anatomical regions comprising auditory core and adjacent auditory-related cortex, a post-hoc ROI analysis was performed. Two different approaches were taken to define the ROIs. We used both a) automatically processed, anatomically defined individual ROIs of Heschl's gyrus (HG) and planum temporale (PT) and b) the well established cytoarchitectonically defined region TE1.0, which is included in the SPM Anatomy toolbox (v1.7; Eickhoff et al., 2005). The TE1.0 region corresponds to the normal location of the core auditory

[1] Note that the functional brain scans covered only about 50 percent of the brain in the inferior-posterior direction, namely the entire perisylvian cortex. For the sake of simplicity, we still refer to this processing step as whole-brain analysis.

cortex on the medial portion of Heschl's gyrus (Morosan et al., 2001; Rademacher et al., 2001). We applied an individual approach, as well as a normalized ROI approach, in order to help generalize our results to different methodologies.

To obtain individual ROIs, the T1 scan was co-registered onto the whole-brain EPI. Subsequently, the whole-brain EPI's normalisation parameters were applied to the T1 scan. This procedure moved it into standard stereotactic space. The normalised anatomical brain scan was then automatically processed with the FreeSurfer software package (v4.5.0; http://surfer.nmr.mgh.harvard.edu; Dale et al., 1999; Fischl, Sereno, & Dale, 1999). After completing the default FreeSurfer processing stream, ROI masks, as provided by FreeSurfer's aparc.a2009s parcellation, were exported individually for each subject's HG and PT. HG includes only the most anterior transverse temporal gyrus. Possible additional transverse temporal gyri are attributed to PT, which comprises both horizontal and vertical aspects (planum parietale Destrieux, Fischl, Dale, & Halgren, 2010). The ROIs only included areas that were fully covered by the functional volumes. Structural overlap maps of the individual ROIs for the entire sample are provided in Supplementary Figure A.1.

Mean beta-values were extracted from first level's contrast images via an in-house tool. This was done for each ROI, COA, and hemisphere for the contrast 80 dB vs. 60 dB. Therefore, reported beta-values represent an increase in effect size from 60 dB to 80 dB. Since we had no interest to explore functional lateralisation, the beta-values for each ROI and COA were averaged across the two hemispheres, so as to improve statistical power. Per ROI, the resulting values were entered into a separate one-way repeated measures ANOVA with COA (7.5, 10, 12.5 and 15 s) as the within-subjects factor for each ROI. Subsequently, linear and quadratic trend analysis was performed on the significant effects.

Standard error, design variance and error variance To obtain information about the first level GLM's error variance (σ^2), mean values within TE1.0 were collected from the ResMS.img of each subject's first level model (Zaehle, Schmidt, et al., 2007). The design variance was calculated from the design matrix and the contrast vector ($c^T(X^TX)^{-1}c$). The standard error was calculated from these measures ($\sigma^2 c^T(X^TX)^{-1}c)^{1/2}$; see Mechelli et al., 2003, Eq. (5)). As with the beta-values, data was averaged across the hemispheres and entered into one-way repeated measures ANOVAs with COA as the within-subjects factor.

To control for the influence of different number of trials per run, all whole-brain and subsequent ROI analyses were performed twice: first entering all collected functional volumes into the model, second entering only the first 30 trials into the model. The first 30 trials of each run were arranged to comprise an equal number of sentences and empty trials (13 60-dB-sentences, 13 80-dB-sentences and 4 empty trials). As a result of this processing step, eight SPM t-tests (80 dB vs. 60 dB: 2 x 4 COA conditions), six ANOVAs at the ROI effect size analyses (2 x HG, PT and TE1.0), two ANOVAs at the analysis of

SEM of parameter estimates, and two ANOVAs at the analysis of error variance will be reported. Since the run with COA = 15 s only contained 30 trials, the all-trials-analysis and the 30-trials-analysis at COA = 15 s are by definition identical.

2.1.4. Results

Behavioural Data

Overall, the evaluation of sentence accent made by research participants was accurate (percent correct (SD): COA = 7.5: 97.4% (2.5), COA = 10: 98.5% (1.9), COA = 12.5: 97.9% (2.9), COA = 15: 98.8% (2.3)). It did not differ between the different COA settings (ANOVA for repeated measures: $F(3,45) = 1.15$, ns).

Whole-brain analysis

For whole-brain second level group analysis, three t-contrasts were calculated for each COA setting separately: an auditory default contrast (all auditory events vs. empty trials) and the two comparisons 80 dB vs. 60 dB and 60 dB vs. 80 dB.

To ensure the general integrity of our analysis, the auditory default contrast was calculated. This revealed bihemispheric clusters for all COA settings in the superior temporal lobe (cf. Supplementary Figure A.2).

Since the comparison of the 80 dB vs. 60 dB contrast over different COA settings indicates differences in the settings' sensitivity in detecting intensity variations in stimuli, this contrast was of major interest. This analysis yielded bihemispheric neuronal activation in the superior temporal plane for each COA setting (FWE, $p < .05$). Figure 2.3 shows the contrast 80 dB vs. 60 dB for all COA settings. The highest amount of significant voxels can be observed at COA = 7.5 s (cf. Table 2.2, Figure 2.4).

To control for the amount of acquired data, additional models were calculated in a second step of analysis. The number of functional brain volumes per COA setting was reduced to the first 30 trials and, as a result, was equal for each COA setting. Once again, the COA = 7.5 s setting revealed the largest clusters. Overall, a lower number of significant voxels can be observed when comparing this analysis to the analysis that comprised all scans.

The reversed t-contrast (60 dB vs. 80 dB) did not result in suprathreshold voxels at any of the COA settings.

ROI analysis

Analyses for the three ROIs were also conducted. Mean beta-values for the 80 dB vs. 60 dB contrast were collected from the individual automatic delineations of HG and PT (created by the FreeSurfer software), as well as the TE1.0 region. Each ROI's mean differential beta-value (80 dB vs. 60 dB) was subjected to a repeated-measures ANOVA, with COA setting as the within-subjects factor. In each ROI a significant main effect

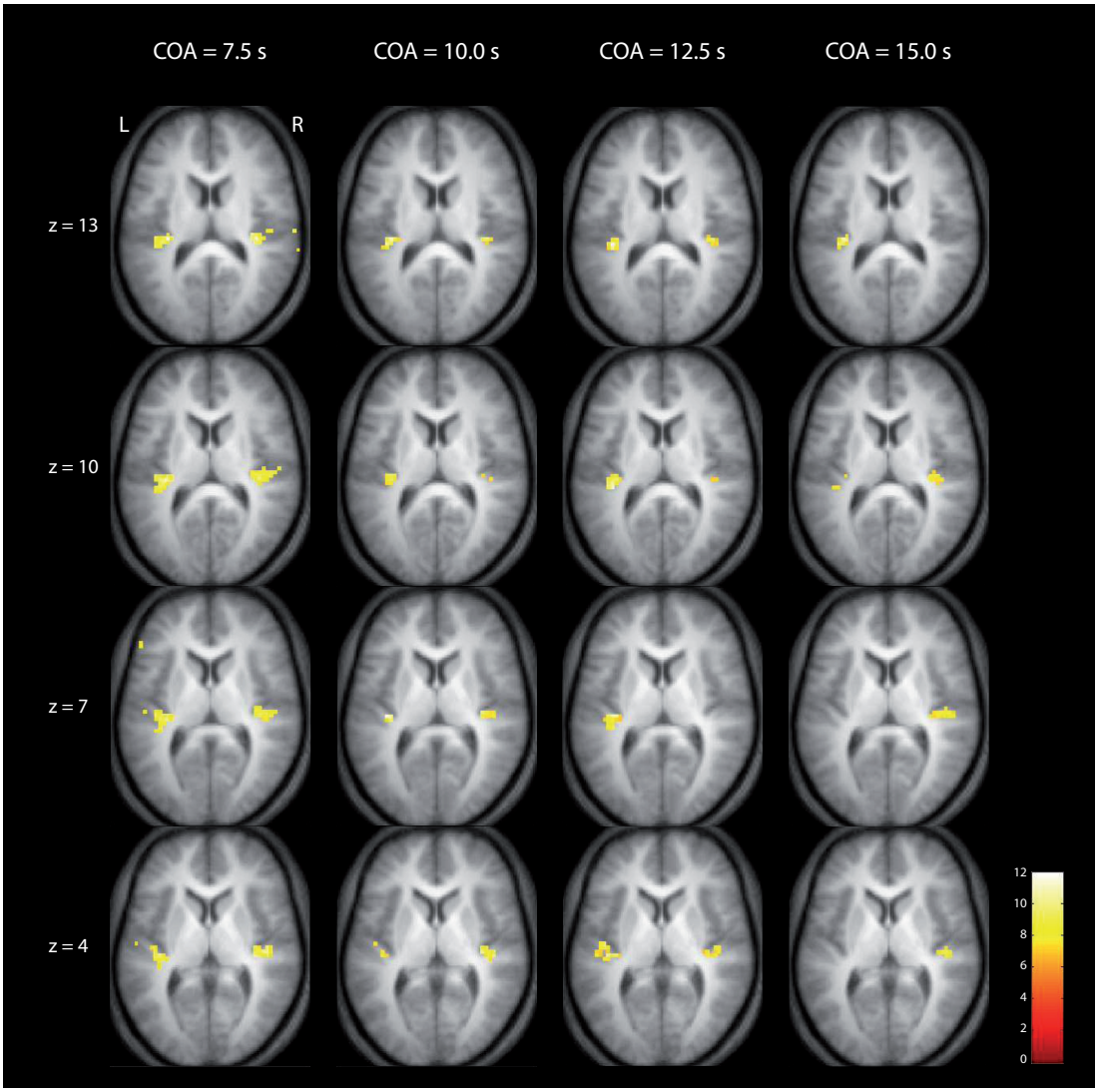

Figure 2.3.: Horizontal slices of the t-contrast 80 dB vs. 60 dB at all four COA settings (all trials) for all subjects (N = 16) projected onto all subjects' mean T1 image (FWE, $p < .05$, $T > 6.8$). MNI space. Neurological convention.

of COA could be found, even when the number of recorded trials was equally balanced (cf. Table 2.3 and Figure 2.5; absolute values are depicted in Supplementary Figure A.3). All linear trends reached statistical significance; in contrast, none of the quadratic trends achieved statistical significance.

In general, the COA = 7.5 s setting resulted in significantly enhanced differential beta-values compared to that of the longer COAs. Therefore, the 7.5 s COA setting results in the largest sensitivity in detecting stimuli differences. This holds true for all three ROIs.

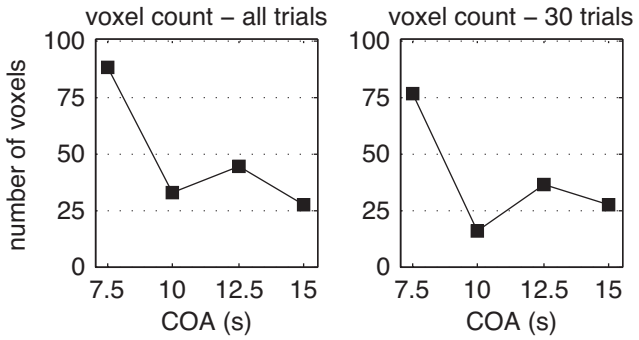

Figure 2.4.: Number of significant voxels in the temporal lobe produced by the second level t-contrast 80 dB vs. 60 dB (FWE, $p < .05$; mean values pooled for both the left and right hemisphere). Left column: all trials. Right column: equal number of trials (30) for each COA setting.

Standard error, design variance and error variance

To elaborate on the models' SEM, design variance and error variance, the design variance was calculated for the contrast 80 dB vs. 60 dB and values for mean error variance were extracted from TE1.0 (auditory core region). This was done only for this ROI, in order to rule out effects of interindividual variability. Then, the standard error was calculated from these measures (cf. Table 2.4). Subsequently, standard error and error variance were subjected to ANOVAs. Values for design variance increased with decreasing number of trials; therefore, increased with increasing COA. The mean error variance increased with increasing COA, though not significantly (when all trials were analysed: $F(1, 19) = 1.4$, ns., Greenhouse-Geisser correction; when an equal number of trials were analyzed: $F(1, 20) = 2.0$, ns., Greenhouse-Geisser correction). SEM of parameter estimates did significantly increase with increasing COA at the all-trials-analysis ($F(1, 20) = 29.2$, $p < .001$, Greenhouse-Geisser correction; post-hoc linear trend: $F(1, 15) = 36.2$, $p < .001$). When equalising the number of trials, the COA settings do not differ significantly with regards to SEM of parameter estimates ($F(1, 22) = 2.3$, ns., Greenhouse-Geisser correction; See Table 2.4).

2.1.5. Discussion

"Silent" scanner protocols are an important tool in auditory fMRI research. Only they allow the presentation of auditory stimuli without disturbance from auditory scanner noise. Standard procedure for sparse and clustered-sparse acquisition schemes is to introduce relatively long silent pauses for stimulus presentation between image acquisitions. It has been previously demonstrated that this approach renders the stimulus-evoked BOLD response largely unaffected by scanner-noise-evoked BOLD response. Therefore, this approach produces data with larger effects (Schmidt et al., 2008). On the downside, relatively long pauses between trials limit the amount of recorded images.

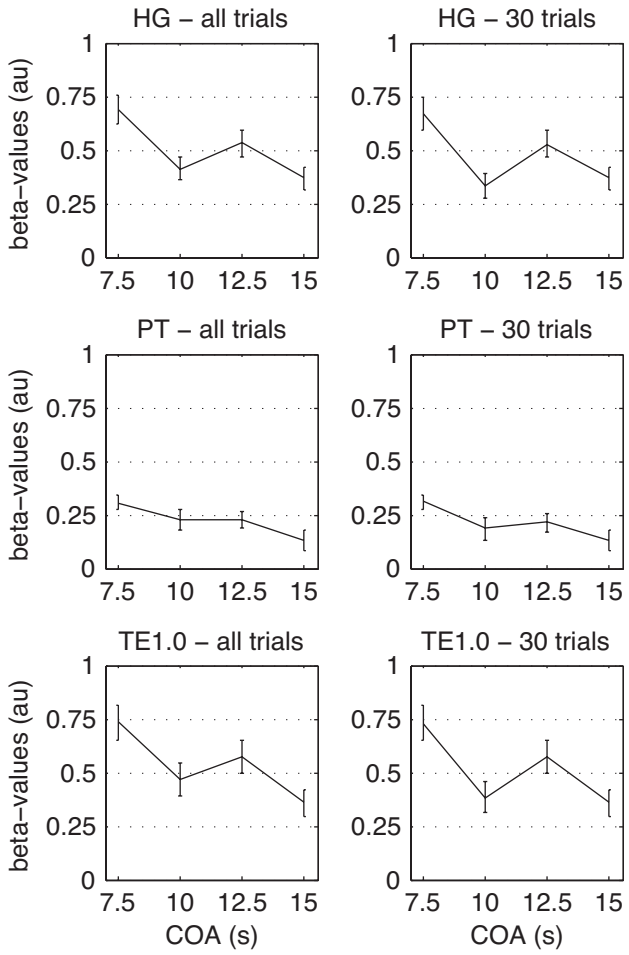

Figure 2.5.: Mean differential beta-values (80 dB vs. 60 dB) for each of the four COAs in the anatomically defined ROIs (± 1 SEM). Left column: all trials. Right column: equal number of trials (30) for each COA setting. HG: Heschl's gyrus; PT: planum temporale; TE1.0: auditory core region. HG and PT ROIs were constructed automatically and individually for each participant (N = 16). All effects are significant (cf. Table 2.3).

This weakens the accuracy of estimated parameters. While the aforementioned reasoning was derived from knowledge about the characteristics of the BOLD response, the aim of this paper is to empirically determine the optimal cluster-onset asynchrony setting for clustered-sparse acquisition schemes in the context of an auditory event-related fMRI design; therefore, determining the optimal duration of the silent pause between image acquisitions. We presented participants with sentences of different intensities in a variety of COA settings, which were implemented in a clustered-sparse acquisition scheme. It has been demonstrated that the clustered-sparse protocol is most advantageous when working with auditory stimuli that consist of a few seconds, presented in an event-related fashion (Zaehle, Schmidt, et al., 2007).

In accordance with present knowledge about the characteristics of the BOLD re-

sponse, we hypothesised that short COA settings would lead to more accurately estimated effects (lower SEM of parameter estimates); this is because short COA settings allow to acquire a higher number of data points. On the other hand, long COA settings, which are presently the standard procedure, should show higher effect sizes (beta-values). Our data confirms the former, but consistently shows that the latter is not the case.

A ROI analysis was performed within auditory core and association areas. Error variance increases with increasing COA, though not significantly. This indicates that at COA = 7.5 s the model is at least as appropriate as at COA = 15 s. Contrary to our expectations, saturation effects might not be a problem at a short COA. Additionally, differences in the subjects' attentional state between different COA settings might influence the characteristics of the BOLD response, which alters error variance. This might vary over different brain regions (Jäncke, Mirzazade, & Shah, 1999; Mechelli et al., 2003; Woods et al., 2009). As both, error variance and the design variance, increase with increasing COA, a short COA setting yielded, as expected, lower SEM of parameter estimates. Nonetheless, this difference vanishes when equalising the amount of acquired data across the COA settings, as predicted by theory and by our hypothesis. Contrary to our expectations, a short COA setting, relative to a long COA setting, led to higher effect sizes (beta-values).

Since the effect sizes at short COAs are larger and more accurately estimated (lower SEM of parameter estimates) than at long COAs, it came as no surprise that a short COA setting yielded a higher number of significant voxels at the whole-brain group statistical analysis. It should be noted that a disproportional large drop in cluster and effect size can be observed at COA = 10 s (cf. Figures 2.4, 2.5). However, the fact that the linear trend analyses gained significance while the quadratic did not suggests a linear decline in beta-values and cluster size. Nevertheless, to draw sound conclusions about the linearity of the decline the number of COA increments needed to be larger. To directly asses the advantages of a short COA in a whole-brain analysis, we calculated the comparison COA = 7.5s vs COA = 15 s (for the 80 dB vs. 60 dB contrast), which yielded bihemispherical suprathreshold voxels in the temporal lobes (cf. Figure 2.6; Note that since the rather conservative voxelwise FWE correction did not yield significant results, we adopted a slightly more liberal approach. We applied a clusterwise FWE $p < .001$ with clusters selected on a voxelwise $p < .001$.). Taken together, these findings demonstrate that a short COA setting (7.5 s) results in higher, more accurately estimated effects. Although not explicitly investigated here, we expect that these results should also be applicable to sparse imaging protocols.

While, error variance and SEM do not differ significantly between the COAs if the number of volume acquisitions is equal, the effect sizes do. This indicates a general advantage of a short COA, irrespective of the amount of acquired data. We conclude that in the context of clustered-sparse designs the hemodynamic response at short COAs (7.5 s in comparison to 15 s) is more dynamically susceptible to external auditory stimuli, at least as far as the auditory-related cortex is concerned. The timing of a short COA setting

Figure 2.6.: Horizontal slices of the t-contrast COA = 7.5 s vs. COA = 15 s (for the comparison 80 dB vs. 60 dB; all trials) for all subjects (N = 16) projected onto all subjects' mean T1 image (clusterwise FWE $p < .001$, clusters selected on a voxelwise $p < .001$). MNI space. Neurological convention

(7.5 s) results in each sentence being presented shortly after the peak of a previous BOLD response (cf. Figure 2.1). One might reason that this occurs during a phase in which a BOLD response can be more easily elicited, than when each sentence was presented later during a phase of possible BOLD signal undershoot.

Investigations of the optimal timing (TR) in auditory fMRI designs have been previously conducted by Shah et al. (2000) and Edmister et al. (1999) in the format of fMRI on/off block designs. Nevertheless, the different stimulus-induced BOLD characteristics in block designs, as compared to event-related clustered-sparse designs, make a comparison between these studies and our approach difficult. Furthermore, as these studies only implemented TRs up to 9 s, no conclusion can be drawn about longer silent pauses.

Compared with continuous fMRI acquisition schemes, sparse, and clustered-sparse designs sample less data in a given period of time. In spite of a less noisy, unobstructed

signal this results in lower statistical power. Thus, improving existing protocols is of utmost importance. By all means, silent protocols are exceptionally beneficial in auditory experiments as they permit the investigation of stimulus perception, which remains unaffected by auditory scanner noise (Schmidt et al., 2008). While it is always of vital interest to reduce sources of error in any measurement tool and to obtain precise data, it becomes even more important when relating two measurements to one another, e.g. behavioural with brain imaging data.

Sentences with a duration of three seconds were used in the present study. Whether or not the present findings also apply to other stimulus durations (e.g. auditory stimuli in the range of hundreds of milliseconds or long speech or music stimuli) might be subject to future studies. Furthermore, the potential influence of attention-related effects might be a topic for further research. For instance, altering the number of stimuli or conditions or changing the task difficulty may alter the participants' attentional state and, as a result, the error variance.

Taken together, these results consistently prove an overall advantage for a short cluster-onset asynchrony setting (7.5 s, i.e. ... 3 s of image acquisition followed by 4.5 s of scanner-silence and stimulus presentation followed by 3 s of image acquisition...) implemented in clustered-sparse fMRI designs. In this study, a short COA resulted in higher differential beta-values and number of activated voxels, irrespective of the number of recorded trials. This opens up the future possibility of increasing the number of trials presented to research participants in a given period of time, therefore increasing the statistical power. Alternatively, an equal number of trials can be delivered in a shorter period. This is especially convenient when working with patients or children because short experiment durations may help make this experience more pleasant and in some cases even possible.

2.1.6. Acknowledgements

This research was supported by the Swiss National Foundation (Grant No. 320030-120661 to MM) and by the "Fonds zur Förderung des akademischen Nachwuchses (FAN) des Zürcher Universitätsvereins (ZUNIV)". We are indebted to Sarah McCourt-Meyer and two anonymous reviewers for helpful comments on an earlier version of this manuscript.

2.1.7. Tables

Table 2.1.: Number of trials for all experimental conditions

COA (s)	**Number of trials per condition**			
	60 dB	80 dB	Empty	Total
7.5	26	26	8	60
10	19	19	7	45
12.5	15	15	6	36
15	13	13	4	30

Table 2.2.: Second level t-test 80 dB vs. 60 dB (FWE, $p < .05$).

COA (s)	**All trials**		**30 trials**	
	T(15)	Voxels	T(15)	Voxels
7.5	10.8	93.5	10.8	77.5
10	10.6	34	8.6	11
12.5	11.9	45	11.6	39.5
15	10.9	30.5	10.9	30.5

T-value of the peak voxel and number of significant voxels in the superior temporal planes. Mean values pooled for both the left and right hemisphere.

Table 2.3.: ROI analysis.

ROI	**All trials**		**30 trials**	
	ANOVA	Linear trend	ANOVA	Linear trend
TE1.0	$F(3, 45) = 8.1$ $p < .001$	$F(1, 15) = 21.2$ $p < .001$	$F(3, 45) = 8.5$ $p < .001$	$F(1, 15) = 17.0$ $p < .01$
HG	$F(3, 45) = 8.0$ $p < .001$	$F(1, 15) = 16.4$ $p < .001$	$F(3, 45) = 7.9$ $p < .001$	$F(1, 15) = 11.2$ $p < .001$
PT	$F(3, 45) = 3.7$ $p < .05$	$F(1, 15) = 9.0$ $p < .01$	$F(3, 45) = 3.1$ $p < .05$	$F(1, 15) = 6.7$ $p < .05$

Main effect COA and linear trend analysis of six separate one-way repeated measures ANOVA.

Table 2.4.: Standard error, design variance and error variance for TE1.0.

COA (s)	Standard error $(\sigma^2 c^T(X^TX)^{-1}c)^{1/2}$	Design variance $c^T(X^TX)^{-1}c$	Error variance σ^2
All trials			
7.5	0.201	0.026	1.638
10	0.245	0.035	1.768
12.5	0.283	0.044	1.920
15	0.310	0.051	2.000
30 trials			
7.5	0.277	0.051	1.542
10	0.294	0.051	1.726
12.5	0.303	0.051	1.911
15	0.309	0.051	1.999

2.2. On the planum temporale lateralization in suprasegmental speech perception. Evidence from a study investigating behavior, structure, and function.

Franziskus Liem, Martina A. Hurschler, Lutz Jäncke, Martin Meyer
Accepted for publication in Human Brain Mapping. doi: 10.1002/hbm.22291

2.2.1. Abstract

This study combines functional and structural magnetic resonance imaging to test the "asymmetric sampling in time" (AST) hypothesis, which makes assertions about the symmetrical and asymmetrical representation of speech in the primary and non-primary auditory cortex. Twenty-three volunteers participated in this parametric clustered-sparse fMRI study. The availability of slowly changing acoustic cues in spoken sentences was systematically reduced over continuous segments with varying lengths (100, 150, 200, 250 ms) by utilizing local time-reversion. As predicted by the hypothesis, functional lateralization in Heschl's gyrus could not be observed. Lateralization in the planum temporale and posterior superior temporal gyrus shifted towards the right hemisphere with decreasing suprasegmental temporal integrity. Cortical thickness of the planum temporale was automatically measured. Participants with an L > R cortical thickness performed better on the in-scanner auditory pattern-matching task. Taken together, these findings support the AST hypothesis and provide substantial novel insight into the division of labor between left and right non-primary auditory cortex functions during comprehension of spoken utterances. In addition, the present data yield support for a structural-behavioral relationship in the non-primary auditory cortex.

2.2.2. Introduction

Long before modern neuroimaging methods have been established, behavioral evidence collected from healthy and brain-injured individuals illustrated that temporal characteristics of speech influence the cortical lateralization during speech perception (Efron, 1963; Schwartz & Tallal, 1980; Studdert-Kennedy & Shankweiler, 1970; Tallal et al., 1996; Zurif & Mendelsohn, 1972). Based on reports of neurological symptoms, namely, pure word deafness, and the psychophysiological properties of speech, Poeppel (2001, 2003) introduced the "asymmetric sampling in time" hypothesis (AST). This framework makes assertions about asymmetrical representation of speech in the auditory-related cortex on the posterior supratemporal plane, which is dependent on the temporal characteristics of the speech signal. It is well established that acoustic information in speech is conveyed over multiple time scales (Rosen, 1992). For instance, the difference between the utterances "date" and "gate" is encoded in rapidly changing formant patterns at the level of acoustic fine structure (subsegmental cues at a scale of several milliseconds up to tens of milliseconds; See also Best, Morrongiello, & Robson, 1981). Conversely, prosodic elements at both the word and sentence levels (e.g. intonation contour, tempo, rhythm, syllabic structure, and stress) are transmitted via slower fluctuations in the speech signal at a scale of hundreds of milliseconds (suprasegmental cues). In functional terms, these parameters convey pivotal information. For instance, the intonation contour (melody) of an utterance can mark the vocalization as either a question or a statement ("That is all?"/"That is all!"). According to the AST hypothesis, the continuously inflowing auditory signal is chunked via temporal integration windows of different sizes. After a suggested initial symmetrical representation of auditory input in core auditory cortex, the acoustic signal is then processed in an asymmetrical manner in the non-primary auditory cortex, presumably the planum temporale (PT). Left non-primary cortex is assumed to predominantly accommodate neuronal ensembles that preferentially integrate over relatively short time windows (up to 50 ms), while right non-primary areas are surmised to harbor neurons that preferentially integrate over relatively long time windows (around 150 - 250 ms) [2]. In terms of speech as an incoming acoustic signal, transient phonetic cues at the subsyllabic level (e.g. formant transitions, voice onset time) preferentially drive neurons residing in the left non-primary auditory cortex, while information that unfolds more slowly (at the syllabic level and beyond, e.g. intonation contour) preferentially drive the right non-primary auditory cortex. The framework is not restricted to speech but also applies to all auditory non-speech signals, for instance music.

[2] Note that the AST framework shares aspects with similar models of auditory processing, most importantly with Zatorre et al.'s temporal/spectral processing model (Zatorre & Belin, 2001; Zatorre, Belin, & Penhune, 2002). These authors propose that left auditory-related cortex is specialized in temporal processing, right in spectral processing. However, the AST hypothesis seems more suited to explain lateralization of speech processing. On the one hand, it incorporates physiologically motivated temporal windows that correspond to linguistic information in the speech code. On the other hand, it avoids categorical dichotomies and instead proposes a continuum of lateralization.

The AST hypothesis has found support from various empirical studies. In a combined EEG-fMRI study, Giraud et al. (2007) have corroborated the notion of endogenous leftward lateralization of fast oscillators (gamma frequencies) and rightward lateralization of slow oscillators (theta frequencies) in the auditory cortex. In general, those oscillations can be seen as foundation of speech processing (Giraud & Poeppel, 2012). Utilizing an array of behavioral and neuroimaging methods, it has been convincingly demonstrated that the processing of rapidly changing acoustic cues results in a right ear advantage and left-lateralized recruitment of auditory-related cortex, respectively (e.g. Belin et al., 1998; Jamison, Watkins, Bishop, & Matthews, 2006; Jäncke et al., 2002; Liégeois-Chauvel, de Graaf, Laguitton, & Chauvel, 1999; Meyer et al., 2005; Obleser, Eisner, & Kotz, 2008; Schönwiesner et al., 2005; Schwartz & Tallal, 1980; Slevc, Martin, Hamilton, & Joanisse, 2011; Studdert-Kennedy & Shankweiler, 1970; Warrier et al., 2009; Zaehle, Wüstenberg, Meyer, & Jäncke, 2004; Zaehle, Jancke, & Meyer, 2007; Zaehle, Jancke, Herrmann, & Meyer, 2009; Zatorre & Belin, 2001; Zatorre et al., 2002). In contrast, evidence for the preference of the right hemisphere for slowly changing acoustic cues is less numerous and less straightforward. Boemio et al. (2005) reported a rightward STS lateralization for slowly fluctuating non-speech stimuli. Meyer et al. found the right PT to be especially sensitive to isolated intonation contour, as well as to speech lacking dynamic pitch variations (Meyer et al., 2002, 2004). These findings correspond with a study by Gandour et al. (2003) that also examined intonation contour. Another suprasegmental domain of speech perception, in which lateralization in posterior temporal regions can be observed, is sentence-level rhyme and rhythm processing. Hurschler, Liem, Jäncke, and Meyer (2012) have shown that an explicit sentence-level rhyming task implicates the right auditory-related cortex more strongly than the left auditory-related cortex. Geiser, Zaehle, Jancke, and Meyer (2008) reported right auditory-related cortex activation during a sentence rhythm detection task. Furthermore, Zhang, Shu, Zhou, Wang, and Li (2010) observed right-lateralized temporal lobe brain activation for processing rhythm and intonation in pseudospeech. Taken together, these results indicate a left hemispheric advantage of auditory-related cortex for the processing of rapidly changing acoustic cues in complex sounds and speech, as well as a right hemispheric preference for slowly changing cues. Nevertheless, due to the present lack of systematic investigations, further studies that examine the processing of slowly changing acoustic cues and the interplay between left and right auditory-related cortex are warranted.

Since the AST hypothesis suggests continuous temporal scales, rather than a temporal dichotomy, the strongest evidence for, or against, the previously postulated predictions might come from a parametric study design. Several studies have utilized complex non-speech sounds in combination with temporal manipulation (e.g. Boemio et al., 2005; Schönwiesner et al., 2005); nevertheless, parametrically designed neuroimaging studies that employ natural speech as stimuli are scarce and are usually limited to very brief stimuli (Britton et al., 2009). However, aforementioned prosodic elements are more pronounced, or only exist in speech exceeding the subsegmental level, for example sen-

tences, as recently emphasized by McGettigan and Scott (2012). Hence, it is of vital interest to test whether the predictions of the AST model are also accurate for longer, more natural sounding, parametrically manipulated speech stimuli.

In order to contribute to a deeper understanding of this issue, the present study examines the effect of temporal integrity of speech input on the functional lateralization in the auditory-related cortex. The current investigation employed an fMRI design, using spoken sentences as stimuli and a stimulus manipulation procedure initially introduced by Saberi and Perrott (1999) and further developed by Walker, Ahmed, and Schnupp (2008). Saberi and Perrott (1999) parametrically manipulated the temporal integrity of speech by segmenting an utterance into consecutive time windows and then by locally time reversing those windows. The longer the length of the locally reversed time segment, the worse the participants performance became; thereby, indicating that the decreasing temporal integrity led to a decrease in comprehensibility. In the present study, four different segment lengths (100 - 250 ms) were utilized, in order to alter the integrity of slow modulations in the speech signal. Additionally, normal sentences (0 ms) were presented as a control condition. First, in order to test the two main predictions of Poeppel's AST model, which posits functional symmetry at the primary auditory cortex, followed by temporally dependent asymmetry at non-primary regions, we investigated the functional lateralization in Heschl's gyrus (HG), PT and the posterior superior temporal gyrus (pSTG) as a function of the stimuli's temporal integrity. We hypothesized that while the lateralization of BOLD response collected from HG remains uninfluenced by the stimuli's temporal integrity, activity in PT and pSTG should shift to the right hemisphere when the availability of intact slowly changing acoustic cues decreases. It has previously been shown that decreasing the amount of acoustic information available in speech stimuli yields increases in activation in auditory-related temporal and frontal regions (Davis & Johnsrude, 2003; Meyer et al., 2004). Therefore, it is reasonable to assume a relatively enhanced signal increase in the right auditory-related cortex when the suprasegmental temporal integrity of stimuli is decreased, resulting in a rightward shift of lateralization in PT and pSTG. Second, the correlation between cortical anatomy, functional and behavioral measurements increasingly gathered attention over the last years. For instance, manual measurements of auditory cortex anatomy were used by Wong et al. (2008) to demonstrate a link between anatomy and success in a pitch learning task, or by Warrier et al. (2009) to show a connection between hemodynamic response and cortical volume. Therefore, we also added an exploratory surface-based morphometry analysis. Here, we inspected the relationship of anatomical PT asymmetry, functional PT lateralization and task performance, in order to explore whether there might be a connection between specific patterns of neuroarchitecture, function and individual differences in task performance during auditory cognition. While the aforementioned studies focused on the relationship between basic acoustic processing and HG anatomy, we chose to focus on the anatomy of the PT, since our task involved more complex speech processing.

2.2.3. Materials and Methods

Participants

Twenty-three healthy native (Swiss-) German speakers took part in this study (12 female, Mage = 25 (SD = 3)). All of the participants were right-handed (Annett, 1970) and had no musical training. Subjects were screened for (developmental) speech and/or hearing impairments, neurological and/or psychiatric history, as well as MR-compatibility. Volunteers gave written informed consent prior to the experiment and were paid for their participation. The Canton of Zurich Ethics Committee approved this study (application E-40/2009).

Stimuli

One hundred and eighty German sentences were binaurally presented in the course of this experiment. They were spoken by a trained female speaker and recorded in a soundproof chamber at the University of Zurich Phonetics Lab. Stimuli were normalized with regards to mean intensity. Each stimulus was presented only once. The temporal structure of the stimuli was manipulated following the approach introduced by Saberi and Perrott (1999). Sentences were split into segments and then locally time reversed. Segment length was varied (100, 150, 200, and 250 ms) to create conditions of differential suprasegmental temporal integrity. Since time reversing segments leads to temporal discontinuities resulting in audible artifacts at the border of two segments, overlapping cosine windows were utilized. This was done via a procedure introduced by Walker et al. (2008) and resulted in artifact free stimuli (cf. Figure 2.7). In addition, unaltered sentences (0 ms) were presented. The sentences' duration was between 2.2 and 3.5 s (M = 2.9 s). The conditions did not differ with regards to sentence duration, stimulus intensity, and intonation contour. Furthermore, probe stimuli were used to implement a pattern-matching task. Each probe stimulus was created from the original, not manipulated version of the stimulus. This was done by cutting an interval with a duration of 10% of the total sentence duration and a random onset. Subsequently, it was assured that the probe stimuli did not contain extended silent periods.

Design and procedure

Participants were familiarized with the stimuli outside the scanner. In addition to the sound attenuating (30 dB) MR headphones (NordicNeuroLab, Bergen, Norway), participants received protective earplugs. The headphones are equipped with high-precision electrostatics transducers and show a flat frequency response between 8 Hz and 35 kHz. Participants were placed inside the scanner and were instructed to keep their eyes open and to look at a fixation cross. Participants saw the cross on a screen via a coil-mounted mirror. Each trial consisted of one sentence, one probe stimulus, and the response phase. The auditory stimuli were presented in a scanner-silent period. The response was given

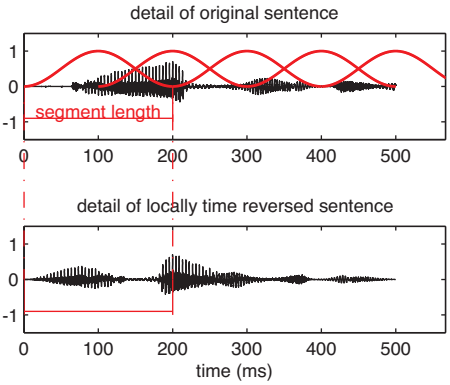

Figure 2.7.: Stimulus manipulation. The top portion illustrates a detail of the waveform of an original sentence, which is multiplied by the cosine windows, and then locally time reversed. In this example, segment length is 200 ms. This procedure results in a bottom waveform.

during the three seconds of image acquisition (See Figure 2.8 for the exact timing)..

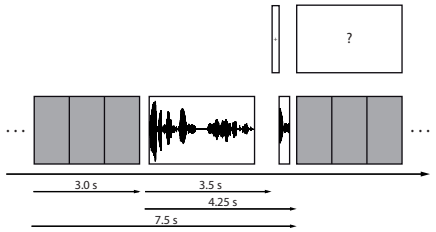

Figure 2.8.: Sequence of a single trial. Gray squares: volume acquisitions; Long waveform: sentence stimulus; Short waveform: probe stimulus; +: fixation cross; ?: response screen

Participants performed a pattern-matching task. They were asked to press a specific button on a button-box to indicate whether or not the probe stimulus was a sample from the original version of the sentence (right index finger for a match response, right middle finger for a no-match response). Sentences were pseudorandomly presented over two runs, balanced across participants. Each of the five conditions (0, 100, 150, 200, 250 ms) was presented 36 times over the course of the experiment. Within each condition, 50% of the probe stimuli were matching fragments. Additionally, 36 empty baseline trials, lacking any sentence and probe stimuli, were presented. During these trials, participants were asked to randomly press a button to control for motor-related brain activity.

MRI data acquisition

Data was acquired at the University Hospital Zurich on a Philips 3T Achieva whole-body MR unit (Philips HealthCare, Best, The Netherlands) equipped with an

eight-channel head coil. Single-shot echo-planar images were recorded using a clustered-sparse acquisition scheme (Liem, Lutz, et al., 2012; Schmidt et al., 2008; Zaehle, Schmidt, et al., 2007). Three volumes were acquired per cluster. In total, 648 functional brain volumes (3 x 216 trials) were collected for each participant (echo planar imaging, 16 transversal slices, in-plane resolution = 2.75 x 2.75 mm, slice thickness = 4 mm, inter-slice gap = 2 mm, matrix size = 80 x 80, field of view [FOV] = 220 x 220 mm, cluster-onset asynchrony [COA] = 7.5 s, acquisition time [TA] = 1 s, echo time [TE] = 35 ms, flip angle = 68°, SENSE factor = 2). When previously using this protocol, we were able to demonstrate that a COA of 7.5 s is more advantageous than standard repetition times/COAs in sparse/clustered-sparse acquisition schemes, which range from 10 s to 15 s (Liem, Lutz, et al., 2012). The functional volumes covered the entire perisylvian and extrasylvian cortex. Additionally, a 3D T1-weighted volume was acquired with a gradient echo sequence (turbo field echo, 160 saggital slices, in-plane resolution = 0.94 x 0.94 mm, slice thickness = 1 mm, matrix size = 256 x 256, FOV = 240 x 240 mm, repetition time [TR] = 8.17 ms, TE = 3.7 ms, flip angle = 8°).

MRI data analysis

The T1-weighted images were analyzed via the Freesurfer software (5.0.0; `http://surfer.nmr.mgh.harvard.edu/`), a) to create individual regions of interest (ROIs) for each subject, which were then employed in the functional ROI analysis and b) to obtain measurements of cortical thickness (CT) and cortical surface area (CSA; Dale & Sereno, 1993; Dale et al., 1999; Fischl, Sereno, & Dale, 1999; Fischl, Sereno, Tootell, & Dale, 1999; Fischl & Dale, 2000; Fischl, Liu, & Dale, 2001; Fischl et al., 2002; Fischl, van der Kouwe, et al., 2004; Fischl, Salat, et al., 2004; Jovicich et al., 2006; Ségonne et al., 2004). In short, Freesurfer reconstructs models of the white matter surface (border between white and gray matter) and the cortical surface (border between gray matter and cerebrospinal fluid). Subsequently, a parcellation of the cortex is performed, based on gyral and sulcal structure (Desikan et al., 2006; Fischl, van der Kouwe, et al., 2004) [3]. Freesurfer provides objective, reliable (Han et al., 2006), and valid (Kuperberg et al., 2003; Rosas et al., 2002; Salat et al., 2004) surface-based measures of cortical anatomy. Each participant's reconstruction was manually checked for accuracy. Measurements of CT and CSA were collected from the a2009s parcellation (Destrieux et al., 2010). Volumetric ROI masks of HG, PT, and pSTG were exported into native space individually for each participant, in order to be used in the context of functional (BOLD related) ROI analysis.

Functional brain volumes were analyzed using SPM8 (`http://www.fil.ion.ucl.ac.uk/spm/software/spm8/`) in a Matlab (R2009a) environment (The MathWorks Inc., MA, USA). We performed a whole brain analysis to check for the general integrity

[3] A more detailed outline of Freesurfer's processing steps is provided by the program developers at `http://surfer.nmr.mgh.harvard.edu/fswiki/FreeSurferWiki?action=AttachFile&do=get&target=freesurfer_methods.doc`.

of our data. For the whole brain analysis, the functional volumes were corrected for head movement and coregistered onto the T1 image. The T1 image was normalized using the unified segmentation approach (Ashburner & Friston, 2005). The resulting normalization matrix was applied to the functional volumes; thereby, transforming them into MNI space (new voxel size = 3 x 3 x 3 mm). Functional images were smoothed with an isotropic 6 mm FWHM Gaussian kernel. A general linear model was applied in which each stimulus was modeled as an event. The hemodynamic response was modeled by means of a boxcar function (Finite Impulse Response, 1st order, 3 s window length). To account for the overall T1 signal decay along the functional volumes of a single cluster, two regressors of no interest were included in the model as established by Zaehle, Schmidt, et al. (2007). Contrast images of the general linear model were entered into a random effects group analysis.

Since we had clear hypotheses with regards to the regions of interest and lateralization, the main results are presented in the ROI analysis. The ROI analysis was performed in an identical manner as the whole brain analysis, with the exception of normalization. To avoid unnecessary interpolation, the ROI analysis was performed in native space. After motion correction, coregistration, smoothing, and modeling the hemodynamic response as mentioned above, mean effect sizes (beta values) within the ROI masks were collected from the model's contrast images for each participant. ROI masks were anatomically defined and created for each subject individually via Freesurfer. We then calculated lateralization indices (LI) for the HG, PT and pSTG ROI (LI = (L - R) / (L + R); possible range of LI: -1 to +1; positive values indicating leftward lateralization, negative values indicating lateralization to the right hemisphere; all entered raw values were positive). The LI values were entered into ANOVAs. Huynh-Feldt correction was used when the assumption of sphericity was violated. When the ANOVAs revealed significant effects, trend analyses (1^{st}, 2^{nd} and 3^{rd} order) were calculated to utilize the parametric study design.

2.2.4. Results

Behavioral results

The in-scanner performance of the pattern-matching task was calculated as the percentage of correct answers per condition. Task performance significantly decreases with increasing segment length (See Figure 2.9; ANOVA for repeated measures, IV: condition, DV: percent correct; $F(3, 64) = 116.53$, $p < .001$, $\eta_p{}^2 = .841$).

MRI results

Whole brain analysis In order to check the general integrity of our data, a whole brain group analysis was performed. For each condition, contrasts against the control condition made up of silent trials were calculated. These contrasts yielded bihemispheric suprathreshold activation in the posterior superior temporal lobe (FWE corrected, p <

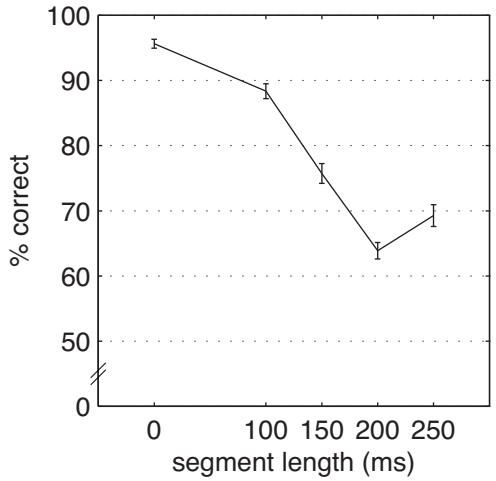

Figure 2.9.: Percent of correct answers from the in-scanner pattern-matching task (± 1 SE).

.05), most notably in the posterior superior temporal lobe. As these contrasts are not of primary interest, they are depicted as supplementary material (Supplementary Figure A.4).

ROI analysis

Functional lateralization of PT and pSTG, but not HG, depends on suprasegmental temporal integrity of stimuli Functional lateralization indices were calculated for HG, PT and pSTG for each condition (for the raw values cf. Table 2.5). While the lateralization in HG is unaffected by the temporal integrity of the stimuli, lateralization in PT and pSTG is significantly influenced. In those two regions, lateralization shifts to the right with increasing segment length. This is demonstrated by a more negative LI (See 2.10; ANOVAs for repeated measures, IV: condition, DV: LI; HG: $F(3, 74) = 2.15$, ns., $\eta_p^2 = .089$; PT: $F(3, 62) = 6.79$, $p < .001$, $\eta_p^2 = .236$, linear trend: $F(1, 22) = 26.44$, $p < .001$, $\eta_p^2 = .546$, higher order trends: ns.; pSTG: $F(2, 52) = 12.6$, $p < 001$, $\eta_p^2 = .365$, linear trend: $F(1, 22) = 17.23$, $p < .001$, $\eta_p^2 = .439$, higher order trends: ns.). This pattern indicates that PT and pSTG are indeed sensitive to the temporal integrity of auditory stimuli.

PT anatomy predicts task performance To explore to what extent the anatomical asymmetry of the PT may influence the performance in the pattern-matching task, LIs for CT_{PT} and CSA_{PT} were calculated (for the raw values cf. Table 2.6). LIs for CT_{PT} and CSA_{PT} did not exhibit sex differences and did not correlate with age or brain volume (CT_{PT}: sex: $t(21) = .16$, ns., age: $r = .07$, ns., intracranial brain volume (ICV): $r = .06$, ns.; CSA_{PT}: sex: $t(21) = -.72$, ns., age: $r = -.05$, ns., ICV: $r = -.06$, ns.). LIs for CT and CSA were not correlated to each other ($r = .355$, ns.). Following Warrier et

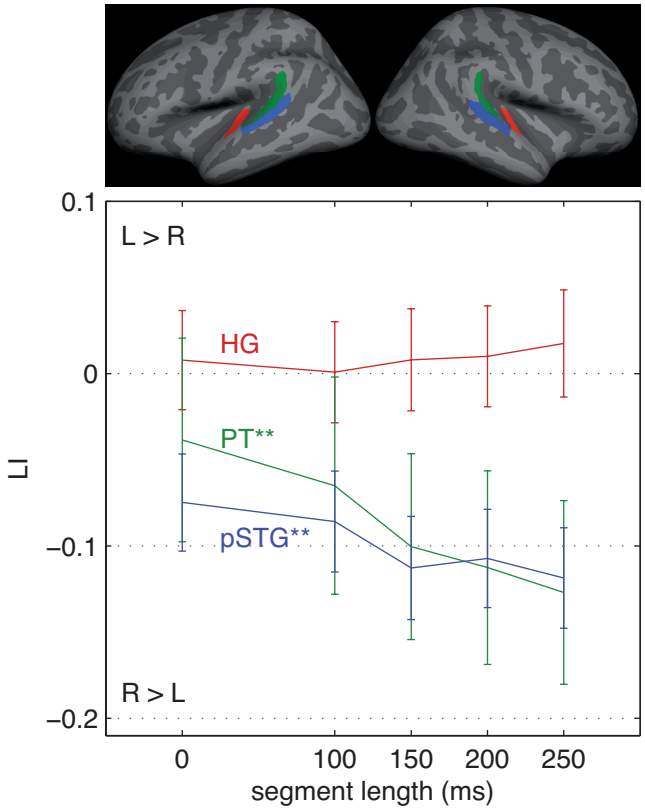

Figure 2.10.: Here, the functional lateralization index (LI; ± 1 SE) in anatomically defined ROIs is displayed; positive values indicate L > R lateralization; HG: Heschl's gyrus (red); PT: planum temporale (green); pSTG: posterior superior temporal gyrus (blue). ** p < .001.

al. (2009), participants were split into two groups: L > R and R > L, for PT thickness and surface area, respectively (for details see Table 2.7). Both groups did not differ with regards to basic characteristics (CT: sex: $\chi^2(1) = .354$, ns., age: t (21) = -.457, ns., ICV: t(21) = .051, ns.; CSA: sex: $\chi^2(1) = .683$, ns., age: t (21) = .667 ns., ICV: t(21) = 1.173, ns.). The thickness-split groups differed significantly with respect to task performance. Participants with L > R thickness perform better than participants with the reverse lateralization pattern (See Figure 2.11; ANOVA, between-subjects IV: $LI_{CT,\ PT}$ group, within-subjects IV: condition, DV: $LI_{CT,\ PT}$; $F_{group}(1, 21) = 5.65$, p < .05, $\eta_p^2 = .212$, $F_{condition}(3, 67) = 111.17$, p < .001, $\eta_p^2 = 0.841$, $F_{group\ X\ condition}(3, 67) = 1.82$, ns., $\eta_p^2 = 0.080$). A closer inspection of the data revealed that the relationship between anatomical asymmetry and behavioral scores is not linear (Pearson correlation between $LI_{CT,\ PT}$ and mean correct answers: r = .26, ns.). Instead, as Supplementary Figure A.5 shows, there seems to exist a categorical difference between R > L and L > R subjects with a lower mean and a larger variability of behavioral performance in the R > L group.

No significant group-related effects could be found for CSA ($F_{group}(1, 21) = 0.07$,

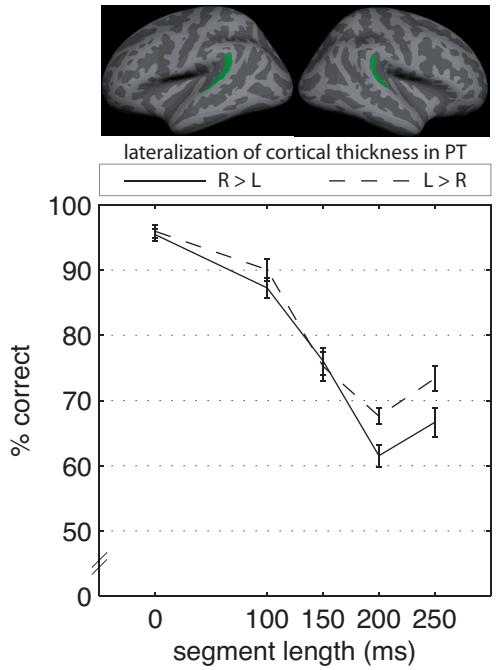

Figure 2.11.: Here, the percent of correct answers from the in-scanner pattern-matching task (± 1 SE) is depicted separately for the two groups of cortical thickness lateralization in planum temporale (PT).

ns., $\eta_p{}^2$= .003, $F_{condition}(3, 62) = 88.25$, $p < .001$, $\eta_p{}^2$= 0.808, $F_{group\ X\ condition}(3, 62) = 1.08$, ns., $\eta_p{}^2$= 0.049).

Relationship between PT function and task performance To assess the relationship between functional lateralization in PT and task performance, we correlated the mean task performance with the functional lateralization in PT for all five conditions separately. This revealed only one significant correlation and two trends towards a significant correlation (r_{0ms} = -.36, $p < .1$; r_{100ms} = -.46, $p < .05$; r_{150ms} = -.25, ns.; r_{200ms} = -.37, $p < .1$; r_{250ms} = -.29, ns.). Here, functional lateralization to the right is associated with better task performance.

Relationship between PT anatomy, function and task performance Finally, we calculated correlations and partial correlations between PT anatomical asymmetry (for cortical thickness and surface area separately), PT functional lateralization (mean of all five conditions) and behavioral performance (mean percent correct). None of the correlations or partial correlations reached statistical significance. However, it has to be noted that CT seems to be more closely related to behavior and function than CSA. Furthermore, the (non-significant) correlation coefficient between cortical thickness asymmetry and functional lateralization was negative; hinting at an inverse relationship between lateralization of structure and function (cf. Supplementary Figure A.6).

2.2.5. Discussion

The present study investigated the AST framework's central hypothesis pertaining to functional lateralization in HG, PT, and pSTG (Poeppel, 2001, 2003). In a parametric event-related clustered-sparse fMRI design, auditory sentences were presented with varying degrees of suprasegmental temporal integrity. Whereas functional lateralization in HG was not present and was actually independent of temporal integrity, a functional shift to the right hemisphere was observed in PT and pSTG when diminishing temporal information with increasing time windows. These findings corroborate the AST framework's notion of a symmetrical representation of the auditory stream in primary auditory cortex and asymmetrical representation in non-primary auditory fields. While left non-primary auditory fields preferentially process rapidly changing acoustic cues, right non-primary fields are supposed to show a preference for slowly changing cues (Poeppel, 2001, 2003). To our knowledge, this is the first study testing the AST's prediction in the context of an fMRI design that employs parametrically manipulated spoken sentences as stimuli. Previous studies using carefully composed sound stimuli (e.g. Boemio et al., 2005; Overath, Kumar, von Kriegstein, & Griffiths, 2008) could not show a clear lateralization in PT or the superior temporal plane. The PT has been termed "computational hub" for complex sound processing that compares the auditory input signal with already stored templates (Griffiths & Warren, 2002). Previous studies have found that the PT is critically involved both in basic auditory and speech processing (Brechmann & Scheich, 2005; Meyer et al., 2002, 2004, 2005; Obleser, Wise, Alex Dresner, & Scott, 2007; Obleser et al., 2008; Xu et al., 2006; Zaehle et al., 2004).

One might be concerned that our results may be driven by a manipulation of "intelligibility", not the manipulation of acoustic cues per se. However, this reasoning contradicts recent evidence that attributed "intelligibility" to the anterior temporal lobe, either in the left or both hemispheres (Friederici, Kotz, Scott, & Obleser, 2010; Obleser, Zimmermann, Van Meter, & Rauschecker, 2007; Scott, Blank, Rosen, & Wise, 2000). Therefore, there is no indication that effects in the regions we investigated (HG, PT, pSTG) were driven by variation of "intelligibility".

We also explored individual differences in cortical anatomy and the relationship to auditory task performance. Recent studies have provided insightful evidence for structural-behavioral relationship in the auditory-related cortex (e.g. Golestani, Molko, Dehaene, LeBihan, & Pallier, 2007; Wong et al., 2008). In the present experiment, lateralization of cortical thickness and surface area in PT was assessed separately as it has been shown that these parameters should be considered independently (Lyttelton et al., 2009). Anatomical PT measurements collected from our sample are well in line with previous reports. A quantitative review by Shapleske, Rossell, Woodruff, and David (1999) showed that an overwhelming majority of studies report a leftward asymmetry of planum temporale surface area. This review shows that, across studies, the mean portion of subjects demonstrating leftward PT asymmetry is around 78 percent. This

fits well with our study where 74 percent of subjects demonstrated leftward surface area asymmetry. Regarding the asymmetry of cortical thickness measures in PT, studies are surprisingly scarce. A study examining post-mortem brains found a rightward lateralization of cortical thickness in the PT (Harasty, Seldon, Chan, Halliday, & Harding, 2003). Around 71 percent of the post-mortem brains demonstrated rightward asymmetry of PT thickness. This is in agreement with our sample where 61 percent of subjects demonstrated a rightward asymmetry. Our results show that while the lateralization of surface area did not predict task performance, the lateralization of cortical thickness did; subjects with L > R lateralization performed better at the auditory pattern-matching task than subjects with R > L lateralization, especially if temporal integrity was severely diminished. It has to be noted that the relationship between the asymmetry of cortical thickness and task performance is not a linear one. It rather seems that there exist a threshold phenomenon, in that rightward CT asymmetry, per se, results in worse task performance.

As cortical thickness in auditory-related cortex tends to be right-lateralized (Harasty et al., 2003) and the left auditory-related cortex exhibits stronger myelination than the right (Sigalovsky, Fischl, & Melcher, 2006), one might argue that subjects with a left-lateralized cortical thickness, thus thinner cortex on the right, might show larger right-hemispheric myelination than subjects with a rightward thickness asymmetry. One might wonder whether this constellation might better support the auditory analysis of suprasegmental cues. Reports of a relationship between cortical thickness in the superior temporal lobes and proficiency in auditory tasks or expertise are relatively scarce (e.g. Bermudez, Lerch, Evans, & Zatorre, 2009; Foster & Zatorre, 2010). Therefore, this topic deserves more attention in future studies.

The correlational analysis of hemodynamic response and task performance yielded a mixed picture. Better task performance tended to be associated with a rightward functional lateralization in the planum temporale. This suggest that a more right-lateralized hemodynamic response better supports suprasegmental analyses required by the task.

Finally, correlation analysis of anatomical asymmetry in PT, functional PT lateralization and task performance did not yield significant results. Nevertheless, this analysis suggests, first, that lateralization of cortical thickness might have a closer relationship to behavior and the lateralization of hemodynamic response than cortical surface area. Second, if anything, lateralization of cortical thickness and hemodynamic response seem to be inversely related. That is, left-lateralized thickness tends to be associated with right-lateralized function. This is in line with a study by (Lu et al., 2009) which found an inverse relationship of cortical thickness and hemodynamic response during an orthographic task. Furthermore, a recent study showed that thinner cortex in auditory-related regions is connected to stronger electrophysiological response during auditory perception (Liem, Zaehle, et al., 2012). The inverse direction of this relationship stands in contrast to the report by Warrier et al. (2009). However, it has to be considered that these au-

thors demonstrated a correlation between cortical volume and extent of the functional response. As we reported cortical thickness and effect sizes of functional response, these results are difficult to compare directly.

2.2.6. Conclusion

In conclusion, we have presented further evidence for the validity of the AST hypothesis in particular with regards to the symmetrical representation of suprasegmental speech in the primary auditory cortex and the asymmetric representation of speech in posterior non-primary auditory cortex. Moreover, we have shown a relationship between asymmetry of planum temporale cortical thickness and behavioral performance during an auditory pattern-matching task.

2.2.7. Acknowledgments

This work was supported by the Swiss National Foundation (grant number 320030-120661 to M.M.) and by the "Fonds zur Förderung des akademischen Nachwuchses (FAN) des Zürcher Universitätsvereins (ZUNIV)". We are indebted to Sarah McCourt Meyer for her helpful comments on this manuscript, Bianca Riesner, Katharina Rufener and Anita Wildi for their help with data acquisition, and Kerry M. M. Walker for sharing the stimulus manipulation procedure.

2.2.8. Tables

Table 2.5.: Raw values for all conditions in the ROIs

Segment length (ms)	**0**		**100**		**150**		**200**		**250**	
Regions	M	SD	M	SD	M	SD	M	SD	M	SD
LH										
HG	5.05	1.35	5.53	1.46	5.56	1.5	5.56	1.45	5.54	1.42
PT	2.54	1.08	2.73	1.15	2.54	1.09	2.61	1.15	2.59	1.14
pSTG	3.67	1.14	3.85	1.18	3.61	1.25	3.73	1.12	3.69	1.14
RH										
HG	5.01	1.38	5.54	1.50	5.5	1.54	5.46	1.41	5.44	1.64
PT	2.82	1.31	3.26	1.43	3.20	1.40	3.29	1.44	3.36	1.39
pSTG	4.24	1.00	4.56	1.12	4.45	1.07	4.57	0.96	4.62	1.00

Raw values (mean beta-values) for all conditions (segment lengths) in the ROIs (HG: Heschl's gyrus, PT: planum temporale, pSTG: posterior superior temporal gyrus). N = 23.

Table 2.6.: Raw values from surface-based morphometry analysis for the planum temporale (CT: cortical thickness, CSA: cortical surface area). N = 23.

CT (mm)				**CSA (mm^2)**			
LH		RH		LH		RH	
M	SD	M	SD	M	SD	M	SD
2.66	0.19	2.69	0.18	675.30	116.58	596.65	139.90

Table 2.7.: Anatomical asymmetry of PT in groups split according to lateralization (CT: cortical thickness, CSA: cortical surface area, LI: lateralization index).

CT						**CSA**					
L > R group			R > L group			L > R group			R > L group		
N	LI	SD	N	LI	SD	N	LI	SD	N	LI	SD
9	0.03	0.02	14	-0.03	0.03	17	0.12	0.08	6	-0.08	0.06

2.3. Cortical thickness of supratemporal plane predicts auditory N1 amplitude.

Franziskus Liem, Tino Zaehle, Anja Burkhard, Lutz Jäncke, Martin Meyer
Published in Neuroreport, 23(17), 1026-30. doi: 10.1097/WNR.0b013e32835abc5c

2.3.1. Abstract

The auditory-evoked potential's N1 component of the scalp electroencephalogram is a well-established measure of electrical brain activity. The N1 reflects basic auditory processing and is modulated by auditory experience, for instance, by musical training. Here, we explore a possible correlation between the auditory N1 amplitude and cortical architecture in the supratemporal plane. We hypothesize that individual differences in N1 amplitude reflect differential acuity, which might also be reflected by differences in auditory cortex anatomy. Auditory potentials evoked by sine wave tones and structural MRI were collected from 27 healthy volunteers. The thickness and surface area of the cortex were calculated using a surface-based morphometry approach. Cortical thickness, rather than surface area, in a cluster on the posterior supratemporal plane, predominantly located on Heschl's sulcus and lateral aspects of Heschl's gyrus, correlated with the N1 amplitude. In particular, lower cortical thickness was associated with larger N1 amplitudes. This is well in agreement with previous functional magnetic resonance studies reporting a thinner cortex to be related to a larger functional response.

2.3.2. Introduction

Auditory-evoked potentials (AEPs) of the scalp electroencephalogram (EEG) are direct measurements of electrical brain activity. The N1 component reflects basic auditory analysis in sound and speech processing (Meyer et al., 2006; Näätänen & Picton, 1987) and remains relatively stable over time within an individual (Roth et al., 1975; Walhovd & Fjell, 2002). The reliability of the N1 amplitude within 12–14 months is up to 0.63 (Walhovd & Fjell, 2002). Furthermore, experience-dependent interindividual differences in the magnitude of the N1 amplitude exist. For instance, several studies have reported differences in AEP components and related measures between musicians and nonmusicians (Baumann et al., 2008; Kühnis, Elmer, Meyer, & Jäncke, 2012; Ott, Langer, Oechslin, Meyer, & Jäncke, 2011). Evidently, Baumann et al. (2008) have reported an enhanced N1 amplitude in musicians during the perception of instrumental and sine wave tones. Among professional musicians, there even occurs an enhanced electromagnetic response when listening to their principal instrument as compared with other instruments (Pantev et al., 2001). Enhanced N1 amplitudes were also observed as a result of speech–sound training (Tremblay & Kraus, 2002). Intracortical recording studies (Godey et al., 2001; Liégeois-Chauvel et al., 1994) as well as magnetoencephalography (MEG) and EEG source estimation approaches (Baumann et al., 2008; Obleser, Scott, & Eulitz, 2006; Yvert et al., 2001; Zaehle, Jancke, & Meyer, 2007) have consistently identified sources in the supratemporal plane responsible for generating the auditory N1. Taken together, the N1 characteristic, similar to neuroanatomical characteristics, varies experience dependently across individuals and, at the same time, is relatively constant within individuals. This suggests a common neuroanatomical foundation within the supratemporal auditory-related cortex.

Examination of cortical morphology in combination with AEPs can shed light on how macroanatomical cortical organization may predict the electrophysiological signal. Although some insight exists on the relationship between functional MRI (fMRI) and structural MRI (sMRI) (Lu et al., 2009), we do not yet know to what extent cortical features (cortical thickness, surface area) may predict the auditory N1 amplitude modulation. Implemented in a multimodal imaging approach, the present study utilized normal individual differences in AEPs and cortical anatomy to explore the relationship between those measures. Therefore, we collected AEP data and structural brain volumes by sMRI from the same individuals. The structural volumes were entered into a surface-based morphometric analysis to correlate auditory N1 amplitude with regional measurements of cortical thickness and surface area in the supratemporal plane. As the P2 component, as compared with the N1, is less reliable and more sensitive to endogenous triggers, such as shifts in attention, we restricted our analysis to the N1 amplitude (Baumann et al., 2008; Harris, Wilson, Eckert, & Dubno, 2012; Roth et al., 1975).

On the basis of the reasoning mentioned above, we hypothesize that the auditory N1 amplitude depends on individual cortical morphology in the human auditory-related

cortex. Intuitively, one would expect a thicker cortex to be related to enhanced N1 amplitudes. However, as there are fMRI studies reporting enhanced brain responses related to a thinner cortex (Lu et al., 2009); this association might be observed in EEG data as well.

2.3.3. Methods

Participants

A total of 27 healthy volunteers (13 women; age: M=25 years, SD=4 years) participated in the study. They were screened for speech and/or hearing impairments, neurological and/or psychiatric history, as well as MR compatibility. All participants were right handed (Annett, 1970) and had no musical training. Volunteers provided written informed consent before the experiment and were paid for their participation. The Canton of Zurich Ethics Committee approved this study (application E-40/2009).

Data acquisition

EEG and MRI data were collected for each participant in two separate sessions. EEG data were recorded using a 128-electrode system (GSN300; Electrical Geodesics Inc., Eugene, Oregon, USA) with Cz as a reference electrode. Impedances were maintained below 30 kΩ. Data were acquired at a sampling frequency of 500 Hz. During the experiment, participants passively listened to 300 binaurally presented stimuli (1 kHz sine wave tones with a duration of 250 ms delivered through in-ear headphones at 70 dB). The interstimulus interval was jittered between 1600 and 2500 ms. The EEG experiment lasted ∼11.5 min. MRI data were collected at the University Hospital Zurich on a Philips 3T Achieva whole-body MR unit (Philips HealthCare, Best, the Netherlands) equipped with an eight-channel head coil. Per participant, one 3D T1-weighted volume was acquired with a gradient echo sequence (turbo field echo, 160 sagittal slices, in-plane resolution=0.94x0.94 mm, slice thickness=1 mm, matrix size=256x256, FOV=240x240 mm, repetition time=8.17 ms, echo time=3.7 ms, flip angle=8°).

Data analysis

EEG data were analyzed using the BrainVision Analyzer software (version 2.0.1; Brain Products, Gilching, Germany). First, data were bandpass filtered (1–30 Hz) and re-referenced onto an average reference. Second, an independent component analysis was carried out to eliminate ocular artifacts (Jung et al., 2000). After interpolating noisy channels, the remaining artifacts were identified and excluded from further analyses. The remaining EEG time series were divided into 500 ms segments (-100 to 400 ms from stimulus onset). Subsequently, these segments were averaged to calculate the AEP at the Cz electrode and baseline corrected (-100 to 0 ms). For each participant, a minimum of 251 trials was entered into the AEP analysis (M=286, SD=12). A grand

average was calculated over the AEPs of all participants (cf. Figure 2.12). The observed grand average N1 peak latency at electrode Cz was then used to define a time window to extract individual N1 peak values (120 ± 20 ms). Cz was chosen in line with the standard protocol, as AEPs have their maximum at frontocentral electrode sites. For the

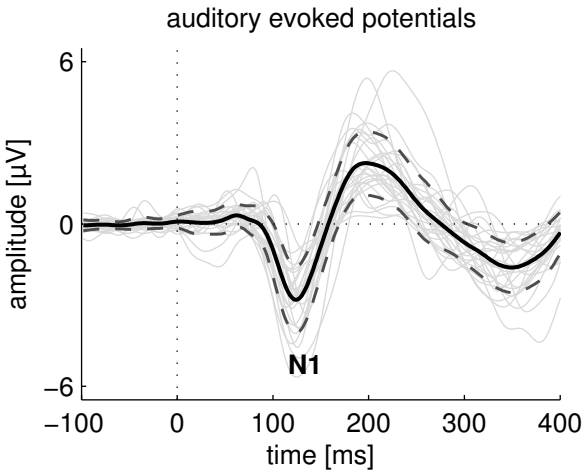

Figure 2.12.: Mean auditory-evoked potential (AEP) waveform at Cz (bold, solid line)±1 SD (bold, dashed lines). Individual waveforms are in gray. Stimulus onset is at 0 ms (N=27).

MRI data analysis, the T1-weighted structural image was processed using the standard Freesurfer processing stream (version 5.1.0; `http://surfer.nmr.mgh.harvard.edu/`). This software offers automated, objective, reliable, and valid algorithms to characterize cortical anatomy (Dale et al., 1999; Fischl, Sereno, & Dale, 1999). Models of the white and pial surface were reconstructed. From these models, vertex-wise values of cortical thickness and surface area were calculated in each participant's native space. For the group analysis, the individual surfaces were registered onto a study-specific average template surface and smoothed with a Gaussian kernel (FWHM of 10 mm). In both hemispheres, statistical analyses were restricted to the supratemporal plane (planum temporale, Heschl's sulcus, Heschl's gyrus, planum polare, and superior temporal gyrus of Freesurfer's a2009s parcellation). N1 peak amplitude values were entered into general linear models to detect correlations between N1 amplitude and the vertex-wise cortical measurements. Subsequently, Monte-Carlo simulations [P(vertex wise)<0.05, P(cluster, corrected)<0.05; n=10 000] were carried out to account for the multiple-comparison problem. This analysis is implemented in the Freesurfer package and assesses the likelihood of significant clusters occurring by chance.

2.3.4. Results

The N1 peak could be identified in all 27 participants. N1 peak amplitude (M=-2.96 μV, SD=1.16 μV, ranging from -5.66 to -1.00 μV; cf. Figure 2.12) did not correlate with participant's age (r=0.28, P=0.15), mean cortical thickness over the whole brain

(r=0.25, P=0.20), or whole-brain surface area (r=0.30, P=0.13). In addition, there were no sex differences in N1 amplitude (t-test for independent samples: t=0.85, P=0.40). Vertex-wise correlation of the N1 peak with cortical thickness yielded one significant cluster on the right supratemporal plane, largely encompassing Heschl's sulcus and lateral aspects of Heschl's gyrus [P(cluster, corrected)<0.001, cluster size=303 mm^2, maximum at MNI coordinates: x=42, y=-29, z=14; cf. Figure 2.13). To check whether the relationship in the left hemisphere is absent at all or only weaker, we carried out an exploratory analysis with a more liberal P threshold. A cluster on the left supratemporal plane could only be observed at P(vertex wise)<0.1 [P(cluster, corrected)=0.015, cluster size=367 mm^2, maximum at MNI coordinates: x=-55, y=-10, z=-3]. Figure 2.14 shows the relationship between individual N1 amplitude and cortical thickness in the significant cluster in the right hemisphere. Figure 2.15 shows the relationship between individual N1 amplitude and cortical thickness in the cluster in the left hemisphere (at the more liberal P threshold). Across participants, a thinner cortex was associated with larger N1 amplitudes. Individual cortical thickness of the right cluster explained 28.3 percent of the interindividual variation in N1 amplitude; individual cortical thickness of the left cluster explained 27.7%. In contrast, vertex-wise correlation of N1 peak with cortical surface area did not yield significant clusters on the supratemporal plane.

2.3.5. Discussion

In the present study, we examined the relationship between the auditory N1 component and two neuroanatomical measurements (cortical thickness and surface area) of the supratemporal plane. The N1 amplitude correlated with cortical thickness in a cluster on the right supratemporal plane, mainly located in Heschl's sulcus and lateral parts of Heschl's gyrus. The thinner the cortex in this cluster, the larger the N1 amplitude. Moreover, the thickness of this auditory-related cortex cluster could explain up to 28.3% of the individual N1 amplitude variation. The location of the cluster is well in agreement with intracranial recordings, which found N1 generators in and around Heschl's gyrus (Godey et al., 2001; Liégeois-Chauvel et al., 1994), as well as with source estimation analyses used in the context of EEG or MEG studies (Baumann et al., 2008; Obleser et al., 2006; Yvert et al., 2001; Zaehle, Jancke, & Meyer, 2007).

To the best of our knowledge, the present study is the first to investigate the relationship between the AEP's N1 component and cortical anatomy using an approach that utilizes normal variations in these measures. A significant relationship was only present for cortical thickness, not for surface area. Even though cortical thickness as well as surface area are partially determined genetically (Winkler et al., 2010), it has been suggested that cortical thickness and surface area are differentially sensitive to environmental influences: although the cortical surface area is considered to be more strongly affected by prenatal events, cortical thickness is apparently more strongly affected by postnatal experience (Frye et al., 2010). Thus, experience-dependent factors

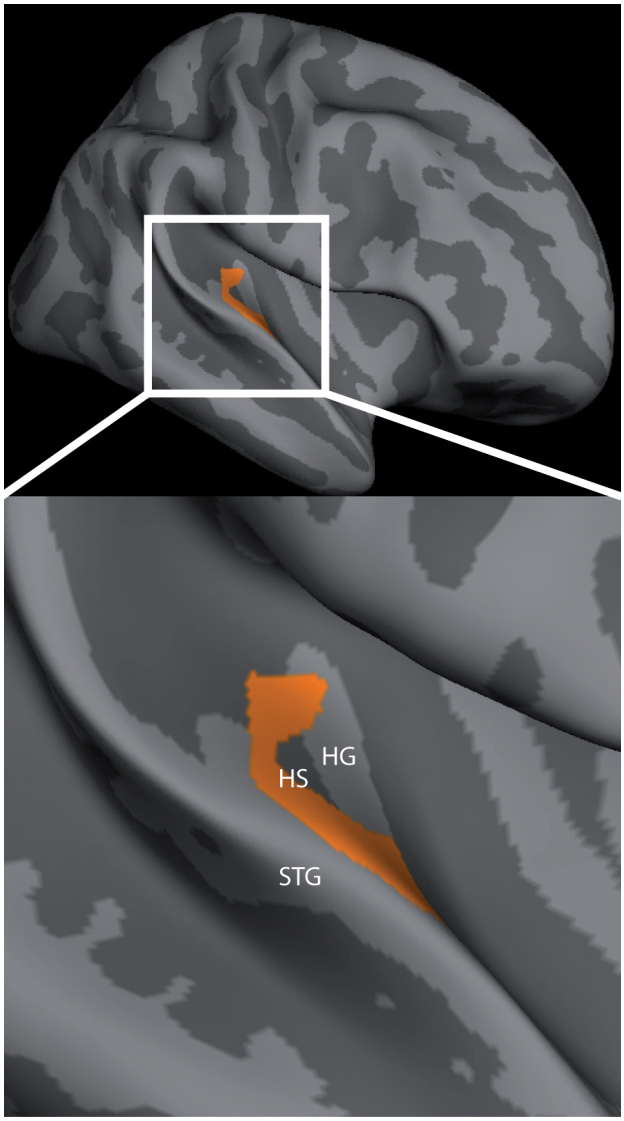

Figure 2.13.: Results of cortical thickness analysis projected on an inflated right hemisphere surface of the average subject template (lateral view) (top). The orange cluster indicates a significant relationship between N1 amplitude and cortical thickness [P(cluster, corrected)<0.001]. Zoom into a significant cluster (bottom). HG, Heschl's gyrus; HS, Heschl's sulcus; STG, superior temporal gyrus.

occurring across the lifespan might alter cortical thickness in a more pronounced manner than cortical surface area. Therefore, the observed variations in cortical thickness and N1 amplitude might be a result of experience-dependent changes because of differences in the auditory environment the individual is raised in and experiences throughout the lifespan. Future longitudinal studies observing the development of auditory proficiency; that is, for instance by examining musicians, at several time points along their musical training, will help shed light on the mechanisms underlying this coupling.

In the present study, a thinner cortex predicted larger N1 amplitudes. This inverse

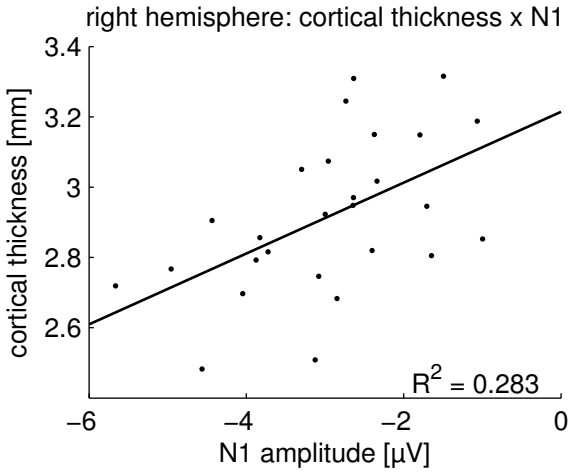

Figure 2.14.: Scatter plot representing the relationship between N1 peak amplitude at Cz and the mean cortical thickness in the cluster on the right supratemporal plane.

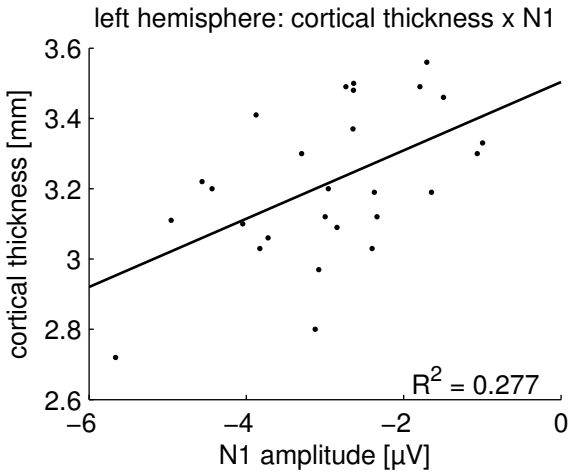

Figure 2.15.: Scatter plot representing the relationship between N1 peak amplitude at Cz and the mean cortical thickness in the cluster on the left supratemporal plane.

relationship of structure and function has been reported previously, in a study combining fMRI and measures of cortical thickness during development (Lu et al., 2009). Stronger activity in frontoparietal regions evoked by an orthographic task was associated with a thinner cortex in this region. There are almost no data on the cellular mechanisms underlying macroscopic changes in gray matter. However, several potential mechanisms are conceivable, for example, changes in the number and morphology of synapses might occur across the lifespan in an experience-dependent manner (Zatorre, Fields, & Johansen-Berg, 2012). Refinement in synaptic circuitry might have resulted in a thinner cortex accommodating neurons that work better synchronized to generate the electrophysiological signal. In general, the simple and appealing idea that a thicker

cortex is better in terms of more efficient neural processing might not hold true in all contexts and brain regions (Hyde et al., 2007), and becomes more and more challenged.

The relationship between N1 amplitude and cortical thickness was statistically stronger in the right supratemporal plane than in the left. Interestingly, there is evidence from an MEG N1 source estimation study that the sources along the right supratemporal plane span a larger area than on the left [25]. In addition, as the stimuli used in our experiment lacked temporal dynamic, our results are also in agreement with the idea that the right auditory-related cortex is specialized in spectral analysis, whereas the left is specialized in temporal analysis (Zatorre et al., 2002). These factors might explain why the relationship between N1 amplitude and cortical thickness was more pronounced in the right hemisphere.

2.3.6. Conclusion

This study provides a first insight into the relationship between cortical architecture and AEPs as revealed by normal individual differences. The findings indicate that the cortical thickness of the supratemporal plane is more strongly related to the AEPs than the cortical surface area. Future studies might help to determine the trajectory-dependent and experience-dependent changes in the relationship between cortical anatomy and AEPs, as well as help to distinguish the influence of genetic and environmental factors on this relationship.

2.3.7. Acknowledgements

This work was supported by the Swiss National Foundation (grant number 320030-120661 to M.M.), by the 'Fonds zur Förderung des akademischen Nachwuchses (FAN) des Zürcher Universitätsvereins (ZUNIV)' and by the German Research Foundation (DFG, SFB/TR 31 to T.Z.). The authors are indebted to James Perry for proofreading this manuscript. They also thank Jürgen Hänggi, Jürg Kühnis, Nicolas Langer, and Cyrill Ott for fruitful discussions.

3. General discussion

In the last part of this thesis, I will (i) discuss the findings yielded by the empirical studies, (ii) elaborate on several open questions, and (iii) sketch recent developments in the field of cognitive neuroscience that likely will result in exciting insight in the future.

3.1. Timing in auditory fMRI

Auditory fMRI experiments are constrained by auditory scanner noise emitted during the switching of gradient coils. In standard fMRI protocols a continuous acquisition is performed. This results in continuous auditory scanner noise, which makes the perception of auditory stimuli difficult and weakens the explanatory power of experiments. In auditory fMRI studies, scanner protocols are used that allow for unhindered stimulus perception by introducing silent pauses between consecutive image acquisitions. These "sparse" or "clustered-sparse" schemes improve auditory fMRI studies with the downside of a reduced amount of data (Gaab et al., 2007a, 2007b; Hall et al., 1999; Schmidt et al., 2008; Zaehle, Schmidt, et al., 2007). Theoretical considerations lead to the conclusion that long silent pauses are beneficial in acquisition, since different hemodynamic responses are separated further in time. As this was never tested empirically in this context, we addressed this issue in Study I (Liem, Lutz, et al., 2012).

Open question 2: Can the theoretically assumed optimal timing for auditory fMRI protocols be supported by empirical evidence?

We hypothesized that (i) long silent pauses result in larger effect sizes, and complementary, (ii) short silent pauses make an effect statistically easier to detect, i.e. increases statistical power due to the increased number of data points. The study aimed to find an equilibrium between those two opposite effects. While (ii) was found to be the case, surprisingly, (i) was irrefutably not. Short pauses yielded preferable results with regards to both, effect size, as well as statistical power. Hence, short silent pauses are recommended in future auditory fMRI schemes.

Due to the nature of the fMRI technique, we can only speculate about neurophysiologic mechanisms that underlie this effect. It appears that in protocols with long silent pauses, stimulus presentation is during a phase in which the hemodynamic response is difficult to elicit. This might be due to BOLD undershoot in the post-peak phase (cf. Figure 2.1).

The findings of Study I enable scientists to perform more advanced auditory fMRI experiments by using short pauses between image acquisitions. Compared to past experiments, they can design studies with increased statistical power if they choose to record an increased amount of data during the same period of time as in past experiments. This might allow them to study smaller samples, which is most important when examining special populations, for instance children or patients, where practical or ethical reasons only allow for a small sample size. Alternatively, they can perform studies with equal statistical power but notably reduced duration, i.e., 50 percent shorter. Time gained by shortening functional experiments might be used to acquire brain images in other modalities, for example additional structural images. Since lying in the confined space of MR scanners while the scanner emits loud, high-pitched sounds is an unpleasant experience for many volunteers, shortening the experiments shortens this unpleasant experience. As a result, data will be less compromised by the volunteer's movement, which gets more pronounced along the course of the experiment. Furthermore, studies with special samples, like children or brain-injured patients, might be simpler to conduct with a short protocol, and, in some cases, even put forth experiments that would be impossible otherwise.

Insights gained from this study were implemented in Study II.

3.2. Functional and structural lateralization in the context of speech processing

Study II addressed hypotheses proposed by the "asymmetric sampling in time" framework suggested by Poeppel (2001, 2003). AST makes assertions about functional lateralization of auditory-related areas as a function of acoustic cues in speech stimuli. While primary auditory cortex (Heschl's gyrus) is supposed to represent the speech signal in a symmetrical manner, non-primary auditory cortex (planum temporale, posterior superior temporal gyrus) is supposed to exhibit cue-dependent functional lateralization; with preferential left-sided lateralization for the processing of rapidly changing acoustic cues and preferential right-sided lateralization for the processing of slowly changing acoustic cues. Previous studies primarily tested this framework via complex, artificial, non-linguistic sound stimuli or very brief linguistic stimuli (syllables). However, since the AST framework aims to explain lateralization of *speech* processing, evidence from experiments using more naturally-sounding and longer stimuli (sentences) is vital. Furthermore, AST makes predictions about lateralization not along a left/right dichotomy, but a left/right continuum. Therefore, a parametrical manipulation along a continuum can yield stronger evidence for the existence of this continuum than a non-parametrical manipulation could.

Recent methodological advances have made it possible to non-invasively study brain structure and to relate structural measurements to behavioral measurements. Cor-

relations between brain structure in auditory-related cortex and performance in auditory tasks have been shown for instance by Golestani et al. (2007) and Wong et al. (2008). Since the planum temporale is fundamentally involved in auditory and speech processing, the additional question arises of whether macroanatomical structural features of this region might predict auditory perception abilities.

Open question 1: Can the functional lateralization predicted by AST be shown with longer, more natural stimuli, in a novel parametric manipulation paradigm? Does anatomical lateralization in planum temporale predict perception in stimuli with altered acoustic information?

To address this question, the present study implemented a previously established behavioral experiment conducted by Saberi and Perrott (1999), in the context of a event-related clustered-sparse fMRI experiment (Liem, Lutz, et al., 2012) in combination with surface-based structural mapping of the planum temporale. For the functional experiment, the temporal integrity of utterances was parametrically manipulated by segmenting a clip into consecutive time windows and then by locally time reversing those windows. The longer the length of the locally reversed time segment, the worse the participants' performance became. As predicted, functional lateralization in Heschl's gyrus was absent, while the planum temporale and the superior temporal gyrus showed a stimulus-dependent pattern of lateralization. Increasing the time windows, thereby decreasing temporal integrity of the stimuli, resulted in a shift of functional lateralization in non-primary auditory regions. Furthermore, automatically measurement of cortical thickness of the planum temporale showed that participants with an L > R cortical thickness performed better on the in-scanner auditory pattern-matching task.

Taken together, these findings support the AST hypothesis and provide substantial novel insight into the division of labor between left and right non-primary auditory cortex function and structure in the context of speech processing. Future studies may further our insight about a broader range of behavioral effects that are based on the observed cortical thickness lateralization (for a further discussion on possible cellular mechanisms that underlie differences in macroanatomy the reader is referred to chapter 3.4).

3.3. Macroanatomical underpinnings of auditory evoked potentials

Study III examined the macroanatomical underpinnings of the auditory evoked potential component N1. For a group of healthy subjects, auditory EEG data as well as anatomical MRI scans were acquired. Via surface-based morphometry, cortical thickness and surface area of the supratemporal plane was measured and correlated with the N1 amplitude. This was done to answer the following question:

Open question 3: Are the characteristics of macroanatomical cortical features in the supratemporal plane a predictor of the auditory N1 amplitude?

The results show that the cortical thickness of a cluster in the supratemporal plane predicts the N1 amplitude measured at the scalp. Interestingly, the direction of the relationship is as follows: the thicker the cortex, the smaller the N1 amplitude. This suggests an efficient organization on neuronal level. Furthermore, since musicianship results in larger N1 amplitudes (Baumann et al., 2008; Pantev et al., 2001), the idea that "thicker cortex is better" might not be true in all contexts (e.g. Hyde et al., 2007); or as Zatorre et al. (2012) wrote:

> "(...) [I]t is not always clear whether training or ability should be associated with increases or decreases in relevant brain regions because of the complex relationship between anatomical changes and underlying functionality." (p. 528)

Regardless of the direction of the effect, the presented electrophysiological-anatomical mapping approach might help to better localize AEPs measured at the scalp. Although sophisticated methods to estimate the source location of scalp potentials have been devised (e.g. Pascual-Marqui, 2002), they can only estimate those sources with centimeter precision. Combining electrophysiology with surface-based morphometry and exploiting normal variations in both measurements might localize the origin of the scalp electrical signal in the brain much more precisely, at the millimeter range.

It has to be noted that little is known about how changes and individual differences in macroanatomy, say cortical thickness and surface area, are evoked on a cellular level.

3.4. Cellular underpinnings of the brain's macroanatomy

We know that brain anatomy is highly variable between healthy humans and can change as a function of learning mechanisms (Zatorre et al., 2012). This is easily assessable with structural MRI and morphometric analysis. However, next to nothing is known about the cellular underpinnings of individual variability and change in macroanatomy. Since investigating the cellular architecture of the brain is only possible with invasive methods, usually via post-mortem brain examination or animal studies, comprehensive basic knowledge in this field is still missing.

In their review article, Zatorre et al. (2012) identify several candidate mechanisms at the cellular level that might result in macroanatomical change. Change in the cerebral cortex, the gray matter, can be due to neuronal or non-neuronal mechanisms. Change in gray matter might arise due to neurogenesis, gliogenesis, synaptogenesis, change in neuronal morphometry or vascular changes. Although the idea was emphatically rejected for a long time, neurogenesis, the birth of new neurons, takes place in the adult mammal

brain (Abrous, Koehl, & Le Moal, 2005). However, whether it is restricted to the hippocampus and the olfactory bulb, or is possible in other areas of the brain as well remains a highly debated topic (Gould, 2007). Other mechanisms seem better equipped to drive cortical changes. Several studies performed on rats reported a task-specific, training-induced creation of new synapses (Kleim et al., 2002) or a change in dendritic morphology (Kolb, Cioe, & Comeau, 2008). These changes are likely to occur in humans, as well. Furthermore, as opposed to the neuronal cells, non-neuronal (glial) cells like astrocytes and oligodendrocytes can divide in the adult cortex. Since gliogenesis is evoked due to learning (Dong & Greenough, 2004), it might also substantially contribute to gray matter changes.

However, one must not ignore non-cellular mechanisms that might be confounded by learning and behavioral changes and also result in changes in the measured MR signal. For example, physical training might alter the vascular structure. In turn, changes in vascular structure might alter apparent macroanatomy without a foundation in neuronal changes.

3.5. Combining multiple brain imaging techniques

In this thesis, a combination of several neuroimaging methods has been used to study brain function and structure. Aggregating multiple multidimensional measures, for instance multiple different structural whole-brain measures, in a single analysis in a way that enables a reasonable interpretation, is a complex issue. The present thesis circumvented this issue by (i) regarding several measures separately (Study II) or (ii) reducing complexity in one measure and correlating this value with a complex, whole-brain measure (Study III). Unavoidably, this implies a loss of information. In recent time, this issue has gathered appreciation in the scientific community. Groves et al. (2012) proposed a "multi-modal fusion analysis" method to integrate data from VBM, SBM and DTI-based measures via a "Linked ICA" (independent component analysis). While this approach reduces dimensionality dramatically, interpretation of the resulting independent components might be difficult and subjective. Since multimodal neuroimaging has gained popularity over the last couple of years, it is highly likely that standard processing pipelines will include similar approaches in the future.

3.6. Future directions – brain networks

As Daugman (2001) noted, the way humans conceptualize the brain and mind is heavily depending on technological advances of the day. Recent advances in communication technology might have facilitated a network view of brain function and structure.

In functional MRI, this is evidenced by the emergence of functional connectivity analysis, especially in the context of so-called "resting state" fMRI. Traditional task-based fMRI studies are interested in characterizing the brain's response to an external

stimulus. Therefore, these studies average time-series that are locked to a stimulus presentation. That way, signal that is unrelated to the external stimulus, i.e. "noise", is reduced. The goal of these studies is to localize a stimulus-related, cognitive function to a brain area. Importantly, changes in brain function evoked by a stimulus can account only for a few percent of the total brain signal. It seemed implausible that the noise, which consumes a majority of the brain's energy, does not contain important information. Therefore, scientists set out to study the "brain's dark energy" (Raichle, 2006); or as Edward Ng, a NASA scientist laconically stated while talking about astrophysical measurements: "one man's noise is another man's signal" (Blakeslee, 1990).

This is done for instance in resting state fMRI studies. Compared to task-based fMRI studies, a very different approach is taken by resting state experiments. Here, no external stimulation is performed. Functional MRI data is acquired over a period of time in which a subject idly lays inside the MR-scanner and tries to think of nothing in particular, while staying awake. Studies conducted in this manner resulted in the discovery of the brain's resting state networks (RSNs; Biswal, Yetkin, Haughton, & Hyde, 1995). Based on signal fluctuations, RSNs constitute systems of regions scattered across the brain that exhibit intrinsic functional coupling. It is assumed that this coupling develops due to previous use, when several regions are repeatedly driven simultaneously by a specific task (Fox & Raichle, 2007).

In the context of speech and language, the conducted resting state studies (e.g. Friederici, Brauer, & Lohmann, 2011; Lohmann et al., 2010; Turken & Dronkers, 2011; Xiang, Fonteijn, Norris, & Hagoort, 2010) do not allow to draw clear conclusions about language/speech RSNs. Hence, future work is urgently required.

In addition to the description of functional networks, structural connectivity can also be studied. This is manly achieved by using diffusion-based MRI techniques (Le Bihan et al., 2001; Le Bihan, 2003), like diffusion tensor imaging (DTI), in combination with diffusion tractography (Jbabdi & Johansen-Berg, 2011). With these methods it became possible to study the brain's structural connections, the white matter, in-vivo. Up until recently, it was only possible to study single tracts, or connections between two or more brain areas. Recent advances in the analysis of complex networks have made it possible to study the connections of the entire brain simultaneously in a relatively straightforward manner (Sporns, Tononi, & Kötter, 2005). Via diffusion imaging and graph theory, the human connectome, the network that represents all connections of the brain, has been described (Hagmann et al., 2008). As these connections substantially influence brain function, studying structure will also illuminate our understanding of brain function (Honey, Thivierge, & Sporns, 2010). Interestingly, Wernicke (1874) already emphasized the importance of the connections between cortical areas in the context of aphasia.

The anatomical connections that support language functions are fairly well known. Those prominent fiber bundles have been mapped via diffusion imaging (Catani, Jones, & ffytche, 2005; Friederici, 2009; Glasser & Rilling, 2008; Rilling, Glasser, Jbabdi, Ander-

sson, & Preuss, 2011; Saur et al., 2008). However, within the connectomic framework it is possible to study the whole brain with the advantage of a variety of different measures that characterize the importance of a single region in the network and, more importantly, the interplay between several regions (Rubinov & Sporns, 2010).

The first studies that examined the relationship of whole-brain white matter networks and language failed to find a relationship between language performance and structural connectivity (Wen et al., 2011). As this study shows several critical flaws, for instance, the crude language tests that were used, future studies with a more refined methodology might provide insight into how language shaped the brain's connectivity. For instance, measures of connectivity might be exploited to characterize the brain's modular organization, i.e. parcellate the cortical anatomy with regards to its connectivity profile (Gorbach et al., 2011). Since this can be achieved at the single subject level, measures of modularity can be correlated with other variables, for instance, test performance.

The connectomic scientific community is characterized by an open and cooperative spirit. Data sharing projects, like the Human Connectome Project (`http://www.humanconnectomeproject.org/`), the "1000 Functional Connectomes" project (`http://www.nitrc.org/projects/fcon_1000/`) or the "Summer of Sharing" (`http://www.childmind.org/en/summer-sharing/`), which share brain imaging data, as well as phenotype data (biographical data, performance in cognitive tests et cetera) open up the possibilities for scientists around the world to collaborate and further advance our understanding of human brain function and structure. Hence, only a computer and analysis skills are required to tackle brain-behavior questions. Since more scientists will be able to conduct research, this will most likely boost the pace of scientific progress. All those projects emphasize the network characteristic of the human brain, indicating a bright future for this concept.

References

Abrous, D. N., Koehl, M., & Le Moal, M. (2005). Adult neurogenesis: from precursors to network and physiology. *Physiol Rev*, *85*(2), 523-69. doi: 10.1152/physrev.00055.2003

Amaro, E., Jr, Williams, S. C. R., Shergill, S. S., Fu, C. H. Y., MacSweeney, M., Picchioni, M. M., ... McGuire, P. K. (2002). Acoustic noise and functional magnetic resonance imaging: current strategies and future prospects. *J Magn Reson Imaging*, *16*(5), 497-510. doi: 10.1002/jmri.10186

Amunts, K., & Zilles, K. (2012). Architecture and organizational principles of Broca's region. *Trends Cogn Sci*, *16*(8), 418-26. doi: 10.1016/j.tics.2012.06.005

Annett, M. (1970). A classification of hand preference by association analysis. *Br J Psychol*, *61*(3), 303-21.

Annett, M. (1992). Five tests of hand skill. *Cortex*, *28*(4), 583-600.

Ashburner, J. (2009). Computational anatomy with the SPM software. *Magn Reson Imaging*, *27*(8), 1163-74. doi: 10.1016/j.mri.2009.01.006

Ashburner, J., & Friston, K. J. (2000). Voxel-based morphometry–the methods. *Neuroimage*, *11*(6 Pt 1), 805-21. doi: 10.1006/nimg.2000.0582

Ashburner, J., & Friston, K. J. (2005). Unified segmentation. *Neuroimage*, *26*(3), 839-51. doi: 10.1016/j.neuroimage.2005.02.018

Bates, E., & Goodman, J. C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. *Language and Cognitive Processes*, *12*(5-6), 507-584. doi: 10.1080/016909697386628

Baumann, S., Meyer, M., & Jäncke, L. (2008). Enhancement of auditory-evoked potentials in musicians reflects an influence of expertise but not selective attention. *J Cogn Neurosci*, *20*(12), 2238-49. doi: 10.1162/jocn.2008.20157

Bear, M. F., Connors, B. W., & Paradiso, M. A. (2007). *Neuroscience: Exploring the brain* (3rd ed.). Lippincott.

Belin, P., Zilbovicius, M., Crozier, S., Thivard, L., Fontaine, A., Masure, M. C., & Samson, Y. (1998). Lateralization of speech and auditory temporal processing. *J Cogn Neurosci*, *10*(4), 536-40.

Ben Shalom, D., & Poeppel, D. (2008). Functional anatomic models of language: assembling the pieces. *Neuroscientist*, *14*(1), 119-27. doi: 10.1177/1073858407305726

Bermudez, P., Lerch, J. P., Evans, A. C., & Zatorre, R. J. (2009). Neuroanatomical correlates of musicianship as revealed by cortical thickness and voxel-based morphometry. *Cereb Cortex*, *19*(7), 1583-96. doi: 10.1093/cercor/bhn196

Best, C. T., Morrongiello, B., & Robson, R. (1981). Perceptual equivalence of acoustic cues in speech and nonspeech perception. *Percept Psychophys*, *29*(3), 191-211.

Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M., & Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *J Neurosci*, *17*(1), 353-62.

Binder, J. R., Frost, J. A., Hammeke, T. A., Rao, S. M., & Cox, R. W. (1996). Function of the left planum temporale in auditory and linguistic processing. *Brain*, *119 (Pt 4)*, 1239-47.

Biswal, B., Yetkin, F. Z., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magn Reson Med*, *34*(4), 537-41.

Blakeslee, S. (1990). Lost on earth: Wealth of data found in space. *The New York Times*, *March 20*.

Boeckx, C. (2009). *Language in cognition*. Wiley-Blackwell. doi: 10.1002/9781444310047

Boemio, A., Fromm, S., Braun, A., & Poeppel, D. (2005). Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nat Neurosci*, *8*(3), 389-95. doi: 10.1038/nn1409

Bonmassar, G., Anami, K., Ives, J., & Belliveau, J. W. (1999). Visual evoked potential (VEP) measured by simultaneous 64-channel EEG and 3T fMRI. *Neuroreport*, *10*(9), 1893-7.

Brandeis, D., Michel, C. M., & Amzica, F. (2009). From neuronal activity to scalp potential fields. In C. M. Michel, T. Koenig, D. Brandeis, L. R. R. Gianotti, & J. Wackermann (Eds.), *Electrical Neuroimaging* (p. 2-24). Cambridge University Press.

Brechmann, A., Baumgart, F., & Scheich, H. (2002). Sound-level-dependent representation of frequency modulations in human auditory cortex: a low-noise fMRI study. *J Neurophysiol*, *87*(1), 423-33.

Brechmann, A., & Scheich, H. (2005). Hemispheric shifts of sound representation in auditory cortex with conceptual listening. *Cereb Cortex*, *15*(5), 578-87. doi: 10.1093/cercor/bhh159

Britton, B., Blumstein, S. E., Myers, E. B., & Grindrod, C. (2009). The role of spectral and durational properties on hemispheric asymmetries in vowel perception. *Neuropsychologia*, *47*(4), 1096-106. doi: 10.1016/j.neuropsychologia.2008.12.033

Broca, P. (1861). Remarque sur le siège de la faculté du langage articulé, suivie d'une observation d'aphemie (perte de la parole). *Bulletin de la Société Anatomique de Paris*, *36*, 330-357.

Buchsbaum, B. R., Olsen, R. K., Koch, P. F., Kohn, P., Kippenhan, J. S., & Berman, K. F. (2005). Reading, hearing, and the planum temporale. *Neuroimage*, *24*(2), 444-54. doi: 10.1016/j.neuroimage.2004.08.025

Catani, M., Jones, D. K., & ffytche, D. H. (2005). Perisylvian language networks of the human brain. *Ann Neurol*, *57*(1), 8-16. doi: 10.1002/ana.20319

Cerasa, A., Morelli, M., Augimeri, A., Salsone, M., Novellino, F., Gioia, M. C., ... Quattrone, A. (2012). Prefrontal thickening in PD with levodopa-induced dyskinesias: New evidence from cortical thickness measurement. *Parkinsonism Relat Disord*. doi: 10.1016/j.parkreldis.2012.06.003

Chalavi, S., Simmons, A., Dijkstra, H., Barker, G. J., & Reinders, A. A. T. S. (2012). Quantitative and qualitative assessment of structural magnetic resonance imaging data in a two-center study. *BMC Med Imaging*, *12*(1), 27. doi: 10.1186/1471-2342-12-27

Clarkson, M. J., Cardoso, M. J., Ridgway, G. R., Modat, M., Leung, K. K., Rohrer, J. D., ... Ourselin, S. (2011). A comparison of voxel and surface based cortical thickness estimation methods. *Neuroimage*, *57*(3), 856-65. doi: 10.1016/j.neuroimage.2011.05.053

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis. I. segmentation and surface reconstruction. *Neuroimage*, *9*(2), 179-94. doi: 10.1006/nimg.1998.0395

Dale, A. M., & Sereno, M. I. (1993). Improved localizadon of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: A linear approach. *Journal of Cognitive Neuroscience*, *5*(2), 162–176. doi: 10.1162/jocn.1993.5.2.162

Daugman, J. (2001). Brain metaphor and brain theory. In W. Bechtel, P. Mandik, J. Mundale, & R. Stufflebeam (Eds.), *Philosophy and the neurosciences.* Blackwell Publishers.

Davis, M. H., & Johnsrude, I. S. (2003). Hierarchical processing in spoken language comprehension. *J Neurosci*, *23*(8), 3423-31.

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., ... Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*, *31*(3), 968-80. doi: 10.1016/j.neuroimage.2006.01.021

Destrieux, C., Fischl, B., Dale, A., & Halgren, E. (2010). Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage*, *53*(1), 1-15. doi: 10.1016/j.neuroimage.2010.06.010

Dong, W. K., & Greenough, W. T. (2004). Plasticity of nonneuronal brain tissue: roles in developmental disorders. *Ment Retard Dev Disabil Res Rev*, *10*(2), 85-90. doi: 10.1002/mrdd.20016

Eden, G. F., Joseph, J. E., Brown, H. E., Brown, C. P., & Zeffiro, T. A. (1999). Utilizing hemodynamic delay and dispersion to detect fMRI signal change without auditory interference: the behavior interleaved gradients technique. *Magn Reson Med*, *41*(1), 13-20.

Edmister, W. B., Talavage, T. M., Ledden, P. J., & Weisskoff, R. M. (1999). Improved auditory cortex imaging using clustered volume acquisitions. *Hum Brain Mapp*, *7*(2), 89-97.

Efron, R. (1963). Temporal perception, aphasia and deja vu. *Brain*, *86*, 403-24.

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., & Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage*, *25*(4), 1325-35. doi: 10.1016/j.neuroimage.2004.12.034

Eling, P. (2011). Lichtheim's golden shot. *Cortex*, *47*(4), 501-8. doi: 10.1016/j.cortex.2010.06.008

Fischl, B., & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proc Natl Acad Sci U S A*, *97*(20), 11050-5. doi: 10.1073/pnas.200033797

Fischl, B., Liu, A., & Dale, A. M. (2001). Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex. *IEEE Trans Med Imaging*, *20*(1), 70-80. doi: 10.1109/42.906426

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., ... Dale, A. M. (2002). Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*, *33*(3), 341-55.

Fischl, B., Salat, D. H., van der Kouwe, A. J. W., Makris, N., Ségonne, F., Quinn, B. T., & Dale, A. M. (2004). Sequence-independent segmentation of magnetic resonance images. *Neuroimage*, *23 Suppl 1*, S69-84. doi: 10.1016/j.neuroimage.2004.07.016

Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis. II: inflation, flattening, and a surface-based coordinate system. *Neuroimage*, *9*(2), 195-207. doi: 10.1006/nimg.1998.0396

Fischl, B., Sereno, M. I., Tootell, R. B., & Dale, A. M. (1999). High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum Brain Mapp*, *8*(4), 272-84.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D. H., ... Dale, A. M. (2004). Automatically parcellating the human cerebral cortex. *Cereb Cortex*, *14*(1), 11-22.

Foster, N. E. V., & Zatorre, R. J. (2010). Cortical structure predicts success in performing musical transformation judgments. *Neuroimage*, *53*(1), 26-36. doi: 10.1016/j.neuroimage.2010.06.042

Fox, M. D., & Raichle, M. E. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci*, *8*(9), 700-11. doi: 10.1038/nrn2201

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends Cogn Sci*, *6*(2), 78-84.

Friederici, A. D. (2009). Pathways to language: fiber tracts in the human brain. *Trends Cogn Sci*, *13*(4), 175-81. doi: 10.1016/j.tics.2009.01.001

Friederici, A. D., Brauer, J., & Lohmann, G. (2011). Maturation of the language network: from inter- to intrahemispheric connectivities. *PLoS One*, *6*(6), e20726. doi: 10.1371/journal.pone.0020726

Friederici, A. D., Kotz, S. A., Scott, S. K., & Obleser, J. (2010). Disentangling syntax and intelligibility in auditory language comprehension. *Hum Brain Mapp*, *31*(3), 448-57. doi: 10.1002/hbm.20878

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., & Frackowiak, R. S. J. (1995). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, *2*(4), 189–210. doi: 10.1002/hbm.460020402

Friston, K. J., Price, C. J., Fletcher, P., Moore, C., Frackowiak, R. S., & Dolan, R. J. (1996). The trouble with cognitive subtraction. *Neuroimage*, *4*(2), 97-104. doi: 10.1006/nimg.1996.0033

Frye, R. E., Liederman, J., Malmberg, B., McLean, J., Strickland, D., & Beauchamp, M. S. (2010). Surface area accounts for the relation of gray matter volume to reading-related skills and history of dyslexia. *Cereb Cortex*, *20*(11), 2625-35. doi: 10.1093/cercor/bhq010

Gaab, N., Gabrieli, J. D. E., & Glover, G. H. (2007a). Assessing the influence of scanner background noise on auditory processing. I. An fMRI study comparing three experimental designs with varying degrees of scanner noise. *Hum Brain Mapp*, *28*(8), 703-20. doi: 10.1002/hbm.20298

Gaab, N., Gabrieli, J. D. E., & Glover, G. H. (2007b). Assessing the influence of scanner background noise on auditory processing. II. An fMRI study comparing auditory processing in the absence and presence of recorded scanner noise using a sparse design. *Hum Brain Mapp*, *28*(8), 721-32. doi: 10.1002/hbm.20299

Galaburda, A. M., Sanides, F., & Geschwind, N. (1978). Human brain. cytoarchitectonic left-right asymmetries in the temporal speech region. *Arch Neurol*, *35*(12), 812-7.

Gandour, J., Dzemidzic, M., Wong, D., Lowe, M., Tong, Y., Hsieh, L., ... Lurito, J. (2003). Temporal integration of speech prosody is shaped by language experience: an fMRI study. *Brain Lang*, *84*(3), 318-36.

Geiser, E., Zaehle, T., Jancke, L., & Meyer, M. (2008). The neural correlate of speech rhythm as evidenced by metrical speech processing. *J Cogn Neurosci*, *20*(3), 541-52. doi: 10.1162/jocn.2008.20029

Geschwind, N. (1965a). Disconnexion syndromes in animals and man. I. *Brain*, *88*(2), 237-94.

Geschwind, N. (1965b). Disconnexion syndromes in animals and man. II. *Brain*, *88*(3), 585-644.

Geschwind, N., & Levitsky, W. (1968). Human brain: left-right asymmetries in temporal speech region. *Science*, *161*(3837), 186-7.

Giraud, A.-L., Kleinschmidt, A., Poeppel, D., Lund, T. E., Frackowiak, R. S. J., & Laufs, H. (2007). Endogenous cortical rhythms determine cerebral specialization for speech perception and production. *Neuron*, *56*(6), 1127-34. doi: 10.1016/j.neuron.2007.09.038

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nat Neurosci*, *15*(4), 511-7. doi: 10.1038/nn.3063

Glasser, M. F., & Rilling, J. K. (2008). DTI tractography of the human brain's language pathways. *Cereb Cortex*, *18*(11), 2471-82. doi: 10.1093/cercor/bhn011

Godey, B., Schwartz, D., de Graaf, J. B., Chauvel, P., & Liégeois-Chauvel, C. (2001). Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: a comparison of data in the same patients. *Clin Neurophysiol*, *112*(10), 1850-9.

Golestani, N., Molko, N., Dehaene, S., LeBihan, D., & Pallier, C. (2007). Brain structure predicts the learning of foreign speech sounds. *Cereb Cortex*, *17*(3), 575-82. doi: 10.1093/cercor/bhk001

Gorbach, N. S., Schütte, C., Melzer, C., Goldau, M., Sujazow, O., Jitsev, J., ... Tittgemeyer, M. (2011). Hierarchical information-based clustering for connectivity-based cortex parcellation. *Front Neuroinform*, *5*, 18. doi: 10.3389/fninf.2011.00018

Gould, E. (2007). How widespread is adult neurogenesis in mammals? *Nat Rev Neurosci*, *8*(6), 481-8. doi: 10.1038/nrn2147

Griffiths, T. D., & Warren, J. D. (2002). The planum temporale as a computational hub. *Trends Neurosci*, *25*(7), 348-53.

Grodzinsky, Y. (2000). The neurology of syntax: language use without Broca's area. *Behav Brain Sci*, *23*(1), 1-21; discussion 21-71.

Grodzinsky, Y., & Santi, A. (2008). The battle for Broca's region. *Trends Cogn Sci*, *12*(12), 474-80. doi: 10.1016/j.tics.2008.09.001

Groves, A. R., Smith, S. M., Fjell, A. M., Tamnes, C. K., Walhovd, K. B., Douaud, G., ... Westlye, L. T. (2012). Benefits of multi-modal fusion analysis on a large-scale dataset: Life-span patterns of inter-subject variability in cortical morphometry and white matter microstructure. *Neuroimage*, *63*(1), 365-380. doi: 10.1016/j.neuroimage .2012.06.038

Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J., & Sporns, O. (2008). Mapping the structural core of human cerebral cortex. *PLoS Biol*, *6*(7), e159. doi: 10.1371/journal.pbio.0060159

Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends Cogn Sci*, *9*(9), 416-23. doi: 10.1016/j.tics.2005.07.004

Hall, D. A., Haggard, M. P., Akeroyd, M. A., Palmer, A. R., Summerfield, A. Q., Elliott, M. R., ... Bowtell, R. W. (1999). "sparse" temporal sampling in auditory fMRI. *Hum Brain Mapp*, *7*(3), 213-23.

Hall, D. A., Summerfield, A. Q., Gonçalves, M. S., Foster, J. R., Palmer, A. R., & Bowtell, R. W. (2000). Time-course of the auditory bold response to scanner noise. *Magn Reson Med*, *43*(4), 601-6.

Han, X., Jovicich, J., Salat, D., van der Kouwe, A., Quinn, B., Czanner, S., ... Fischl, B. (2006). Reliability of MRI-derived measurements of human cerebral cortical thickness: the effects of field strength, scanner upgrade and manufacturer. *Neuroimage*, *32*(1), 180-94. doi: 10.1016/j.neuroimage.2006.02.051

Harasty, J., Seldon, H. L., Chan, P., Halliday, G., & Harding, A. (2003). The left human speech-processing cortex is thinner but longer than the right. *Laterality*, *8*(3), 247-60. doi: 10.1080/13576500244000175

Harris, K. C., Wilson, S., Eckert, M. A., & Dubno, J. R. (2012). Human evoked cortical activity to silent gaps in noise: effects of age, attention, and cortical processing speed. *Ear Hear*, *33*(3), 330-9. doi: 10.1097/AUD.0b013e31823fb585

Hart, H. C., Hall, D. A., & Palmer, A. R. (2003). The sound-level-dependent growth in the extent of fMRI activation in Heschl's gyrus is different for low- and high-frequency tones. *Hear Res*, *179*(1-2), 104-12.

Herrmann, C. S., & Debener, S. (2008). Simultaneous recording of EEG and BOLD responses: a historical perspective. *Int J Psychophysiol*, *67*(3), 161-8. doi: 10.1016/j.ijpsycho.2007.06.006

Herrmann, C. S., Oertel, U., Wang, Y., Maess, B., & Friederici, A. D. (2000). Noise affects auditory and linguistic processing differently: an MEG study. *Neuroreport*, *11*(2), 227-9.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nat Rev Neurosci*, *8*(5), 393-402. doi: 10.1038/nrn2113

Hinshaw, W. S., Bottomley, P. A., & Holland, G. N. (1977). Radiographic thin-section image of the human wrist by nuclear magnetic resonance. *Nature*, *270*(5639), 722-3.

Honey, C. J., Thivierge, J.-P., & Sporns, O. (2010). Can structure predict function in the human brain? *Neuroimage*, *52*(3), 766-76. doi: 10.1016/j.neuroimage.2010.01.071

Hu, S., Olulade, O., Castillo, J. G., Santos, J., Kim, S., Tamer, G. G., Jr, ... Talavage, T. M. (2010). Modeling hemodynamic responses in auditory cortex at 1.5 T using variable duration imaging acoustic noise. *Neuroimage*, *49*(4), 3027-38. doi: 10.1016/j.neuroimage.2009.11.051

Hurschler, M. A., Liem, F., Jäncke, L., & Meyer, M. (2012). Right and left perisylvian cortex and left inferior frontal cortex mediate sentence-level rhyme detection in spoken language as revealed by sparse fMRI. *Hum Brain Mapp*. doi: 10.1002/hbm.22134

Hyde, K. L., Lerch, J. P., Zatorre, R. J., Griffiths, T. D., Evans, A. C., & Peretz, I. (2007). Cortical thickness in congenital amusia: when less is better than more. *J Neurosci*, *27*(47), 13028-32. doi: 10.1523/JNEUROSCI.3039-07.2007

Schwarzbauer, C., Davis, M. H., Rodd, J. M., & Johnsrude, I. (2006). Interleaved silent steady state (ISSS) imaging: a new sparse imaging method applied to auditory fmri. *Neuroimage*, *29*(3), 774-82. doi: 10.1016/j.neuroimage.2005.08.025

Jamison, H. L., Watkins, K. E., Bishop, D. V. M., & Matthews, P. M. (2006). Hemispheric specialization for processing auditory nonspeech stimuli. *Cereb Cortex*, *16*(9), 1266-75. doi: 10.1093/cercor/bhj068

Jäncke, L. (2005). *Methoden der Bildgebung in der Psychologie und den kognitiven Neurowissenschaften*. Kohlhammer.

Jäncke, L., Mirzazade, S., & Shah, N. J. (1999). Attention modulates activity in the primary and the secondary auditory cortex: a functional magnetic resonance imaging study in human subjects. *Neurosci Lett*, *266*(2), 125-8.

Jäncke, L., Shah, N. J., Posse, S., Grosse-Ryuken, M., & Müller-Gärtner, H. W. (1998). Intensity coding of auditory stimuli: an fMRI study. *Neuropsychologia*, *36*(9), 875-83.

Jäncke, L., Wüstenberg, T., Scheich, H., & Heinze, H.-J. (2002). Phonetic perception and the temporal cortex. *Neuroimage*, *15*(4), 733-46. doi: 10.1006/nimg.2001.1027

Jbabdi, S., & Johansen-Berg, H. (2011). Tractography: where do we go from here? *Brain Connect*, *1*(3), 169-83. doi: 10.1089/brain.2011.0033

Jovicich, J., Czanner, S., Greve, D., Haley, E., van der Kouwe, A., Gollub, R., ... Dale, A. (2006). Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data. *Neuroimage*, *30*(2), 436-43. doi: 10.1016/j.neuroimage.2005.09.046

Jung, T. P., Makeig, S., Humphries, C., Lee, T. W., McKeown, M. J., Iragui, V., & Sejnowski, T. J. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, *37*(2), 163-78.

Kleim, J. A., Barbay, S., Cooper, N. R., Hogg, T. M., Reidel, C. N., Remple, M. S., & Nudo, R. J. (2002). Motor learning-dependent synaptogenesis is localized to functionally reorganized motor cortex. *Neurobiol Learn Mem*, *77*(1), 63-77. doi: 10.1006/nlme.2000.4004

Kolb, B., Cioe, J., & Comeau, W. (2008). Contrasting effects of motor and visual spatial learning tasks on dendritic arborization and spine density in rats. *Neurobiol Learn Mem*, *90*(2), 295-300. doi: 10.1016/j.nlm.2008.04.012

Kühnis, J., Elmer, S., Meyer, M., & Jäncke, L. (2012). Musicianship boosts perceptual learning of pseudoword-chimeras: An electrophysiological approach. *Brain Topogr*. doi: 10.1007/s10548-012-0237-y

Kuperberg, G. R., Broome, M. R., McGuire, P. K., David, A. S., Eddy, M., Ozawa, F., ... Fischl, B. (2003). Regionally localized thinning of the cerebral cortex in schizophrenia. *Arch Gen Psychiatry*, *60*(9), 878-88. doi: 10.1001/archpsyc.60.9.878

Le Bihan, D. (2003). Looking into the functional architecture of the brain with diffusion MRI. *Nat Rev Neurosci*, *4*(6), 469-80. doi: 10.1038/nrn1119

Le Bihan, D., Mangin, J. F., Poupon, C., Clark, C. A., Pappata, S., Molko, N., & Chabriat, H. (2001). Diffusion tensor imaging: concepts and applications. *J Magn Reson Imaging*, *13*(4), 534-46.

Li, J., Li, W., Xian, J., Li, Y., Liu, Z., Liu, S., ... He, H. (2012). Cortical thickness analysis and optimized voxel-based morphometry in children and adolescents with prelingually profound sensorineural hearing loss. *Brain Res*, *1430*, 35-42. doi: 10.1016/j.brainres.2011.09.057

Lichtheim, L. (1885). On aphasia. *Brain*, *7*, 433-484.

Liégeois-Chauvel, C., de Graaf, J. B., Laguitton, V., & Chauvel, P. (1999). Specialization of left auditory cortex for speech perception in man depends on temporal coding. *Cereb Cortex*, *9*(5), 484-96.

Liégeois-Chauvel, C., Musolino, A., Badier, J. M., Marquis, P., & Chauvel, P. (1994). Evoked potentials recorded from the auditory cortex in man: evaluation and topography of the middle latency components. *Electroencephalogr Clin Neurophysiol*, *92*(3), 204-14.

Liem, F., Hurschler, M. A., Jäncke, L., & Meyer, M. (2013). On the planum temporale lateralization in suprasegmental speech perception: Evidence from a study investigating behavior, structure, and function. *Human Brain Mapping*. doi: 10.1002/hbm.22291

Liem, F., Lutz, K., Luechinger, R., Jäncke, L., & Meyer, M. (2012). Reducing the interval between volume acquisitions improves "sparse" scanning protocols in event-related auditory fMRI. *Brain Topogr*, *25*(2), 182-93. doi: 10.1007/s10548-011-0206-x

Liem, F., Zaehle, T., Burkhard, A., Jäncke, L., & Meyer, M. (2012). Cortical thickness of supratemporal plane predicts auditory N1 amplitude. *Neuroreport*, *23*(17), 1026-30. doi: 10.1097/WNR.0b013e32835abc5c

Lohmann, G., Hoehl, S., Brauer, J., Danielmeier, C., Bornkessel-Schlesewsky, I., Bahlmann, J., ... Friederici, A. (2010). Setting the frame: the human brain activates a basic low-frequency network for language processing. *Cereb Cortex*, *20*(6), 1286-92. doi: 10.1093/cercor/bhp190

Lu, L. H., Dapretto, M., O'Hare, E. D., Kan, E., McCourt, S. T., Thompson, P. M., ... Sowell, E. R. (2009, Nov). Relationships between brain activation and brain structure in normally developing children. *Cereb Cortex*, *19*(11), 2595-604. doi: 10.1093/cercor/bhp011

Lyttelton, O. C., Karama, S., Ad-Dab'bagh, Y., Zatorre, R. J., Carbonell, F., Worsley, K., & Evans, A. C. (2009). Positional and surface area asymmetry of the human cerebral cortex. *Neuroimage*, *46*(4), 895-903. doi: 10.1016/j.neuroimage.2009.03.063

May, A., & Gaser, C. (2006). Magnetic resonance-based morphometry: a window into structural plasticity of the brain. *Curr Opin Neurol*, *19*(4), 407-11. doi: 10.1097/01.wco.0000236622.91495.21

McGettigan, C., & Scott, S. K. (2012). Cortical asymmetries in speech perception: what's wrong, what's right and what's left? *Trends Cogn Sci*, *16*(5), 269-76. doi: 10.1016/j.tics.2012.04.006

Mechelli, A., Price, C. J., Henson, R. N. A., & Friston, K. J. (2003). Estimating efficiency a priori: a comparison of blocked and randomized designs. *Neuroimage*, *18*(3), 798-805.

Meyer, M., Alter, K., Friederici, A. D., Lohmann, G., & von Cramon, D. Y. (2002). FMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. *Hum Brain Mapp*, *17*(2), 73-88. doi: 10.1002/hbm.10042

Meyer, M., Baumann, S., & Jancke, L. (2006). Electrical brain imaging reveals spatio-temporal dynamics of timbre perception in humans. *Neuroimage*, *32*(4), 1510-23. doi: 10.1016/j.neuroimage.2006.04.193

Meyer, M., Steinhauer, K., Alter, K., Friederici, A. D., & von Cramon, D. Y. (2004). Brain activity varies with modulation of dynamic pitch variance in sentence melody. *Brain Lang*, *89*(2), 277-89. doi: 10.1016/S0093-934X(03)00350-X

Meyer, M., Zaehle, T., Gountouna, V.-E., Barron, A., Jancke, L., & Turk, A. (2005). Spectro-temporal processing during speech perception involves left posterior auditory cortex. *Neuroreport*, *16*(18), 1985-9.

Mietchen, D., & Gaser, C. (2009). Computational morphometry for detecting changes in brain structure due to development, aging, learning, disease and evolution. *Front Neuroinform*, *3*, 25. doi: 10.3389/neuro.11.025.2009

Moelker, A., & Pattynama, P. M. T. (2003). Acoustic noise concerns in functional magnetic resonance imaging. *Hum Brain Mapp*, *20*(3), 123-41. doi: 10.1002/hbm .10134

Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., & Zilles, K. (2001). Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system. *Neuroimage*, *13*(4), 684-701. doi: 10.1006/nimg.2000 .0715

Mosso, A. (1881). *Ueber Den Kreislauf Des Blutes Im Menschlichen Gehirn*. Reprinted by Books on Demand, 2010.

Mueller, K., Mildner, T., Fritz, T., Lepsien, J., Schwarzbauer, C., Schroeter, M. L., & Möller, H. E. (2011). Investigating brain response to music: a comparison of different fMRI acquisition schemes. *Neuroimage*, *54*(1), 337-43. doi: 10.1016/j.neuroimage.2010 .08.029

Mulert, C., Jäger, L., Propp, S., Karch, S., Störmann, S., Pogarell, O., ... Hegerl, U. (2005). Sound level dependence of the primary auditory cortex: Simultaneous measurement with 61-channel EEG and fMRI. *Neuroimage*, *28*(1), 49-58. doi: 10.1016/ j.neuroimage.2005.05.041

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology*, *24*(4), 375-425.

Obleser, J., Eisner, F., & Kotz, S. A. (2008). Bilateral speech comprehension reflects differential sensitivity to spectral and temporal features. *J Neurosci*, *28*(32), 8116-23. doi: 10.1523/JNEUROSCI.1290-08.2008

Obleser, J., Scott, S. K., & Eulitz, C. (2006). Now you hear it, now you don't: transient traces of consonants and their nonspeech analogues in the human brain. *Cereb Cortex*, *16*(8), 1069-76. doi: 10.1093/cercor/bhj047

Obleser, J., Wise, R. J. S., Alex Dresner, M., & Scott, S. K. (2007). Functional integration across brain regions improves speech perception under adverse listening conditions. *J Neurosci*, *27*(9), 2283-9. doi: 10.1523/JNEUROSCI.4663-06.2007

Obleser, J., Zimmermann, J., Van Meter, J., & Rauschecker, J. P. (2007). Multiple stages of auditory speech perception reflected in event-related fMRI. *Cereb Cortex*, *17*(10), 2251-7. doi: 10.1093/cercor/bhl133

Ogawa, S., Lee, T. M., Nayak, A. S., & Glynn, P. (1990). Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. *Magn Reson Med*, *14*(1), 68-78.

Ott, C. G. M., Langer, N., Oechslin, M. S., Meyer, M., & Jäncke, L. (2011). Processing of voiced and unvoiced acoustic stimuli in musicians. *Front Psychol*, *2*, 195. doi: 10.3389/fpsyg.2011.00195

Overath, T., Kumar, S., von Kriegstein, K., & Griffiths, T. D. (2008). Encoding of spectral correlation over time in auditory cortex. *J Neurosci*, *28*(49), 13268-73. doi: 10.1523/JNEUROSCI.4596-08.2008

Panizzon, M. S., Fennema-Notestine, C., Eyler, L. T., Jernigan, T. L., Prom-Wormley, E., Neale, M., ... Kremen, W. S. (2009). Distinct genetic influences on cortical surface area and cortical thickness. *Cereb Cortex*, *19*(11), 2728-35. doi: 10.1093/cercor/bhp026

Pantev, C., Roberts, L. E., Schulz, M., Engelien, A., & Ross, B. (2001). Timbre-specific enhancement of auditory cortical representations in musicians. *Neuroreport*, *12*(1), 169-74.

Pascual-Marqui, R. D. (2002). Standardized low-resolution brain electromagnetic tomography (sloreta): technical details. *Methods Find Exp Clin Pharmacol*, *24 Suppl D*, 5-12.

Pauling, L., & Coryell, C. D. (1936). The magnetic properties and structure of hemoglobin, oxyhemoglobin and carbonmonoxyhemoglobin. *Proc Natl Acad Sci U S A*, *22*(4), 210-6.

Pereira, J. B., Ibarretxe-Bilbao, N., Marti, M.-J., Compta, Y., Junqué, C., Bargallo, N., & Tolosa, E. (2011). Assessment of cortical degeneration in patients with Parkinson's disease by voxel-based morphometry, cortical folding, and cortical thickness. *Hum Brain Mapp*. doi: 10.1002/hbm.21378

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. E. (1988). Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature*, *331*(6157), 585-9. doi: 10.1038/331585a0

Poeppel, D. (2001). Pure word deafness and the bilateral processing of the speech code. *Cognitive Science*, *25*(5), 679 - 693. doi: 10.1016/S0364-0213(01)00050-7

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. *Speech Communication*, *41*(1), 245 - 255. doi: 10.1016/S0167-6393(02)00107-3

Poeppel, D., & Embick, D. (2005). Defining the Relation Between Linguistics and Neuroscience. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones* (p. 103-116). Lawrence Erlbaum Assoc.

Poeppel, D., & Hickok, G. (2004). Towards a new functional anatomy of language. *Cognition*, *92*(1-2), 1-12. doi: 10.1016/j.cognition.2003.11.001

Pulvermüller, F. (2010). Brain embodiment of syntax and grammar: discrete combinatorial mechanisms spelt out in neuronal circuits. *Brain Lang*, *112*(3), 167-79. doi: 10.1016/j.bandl.2009.08.002

Rademacher, J., Morosan, P., Schormann, T., Schleicher, A., Werner, C., Freund, H. J., & Zilles, K. (2001). Probabilistic mapping and volume measurement of human primary auditory cortex. *Neuroimage*, *13*(4), 669-83. doi: 10.1006/nimg.2000.0714

Raichle, M. E. (1983). Positron emission tomography. *Annu Rev Neurosci*, *6*, 249-67. doi: 10.1146/annurev.ne.06.030183.001341

Raichle, M. E. (2006). Neuroscience. the brain's dark energy. *Science*, *314*(5803), 1249-50. doi: 10.1126/science.1134405

Raichle, M. E. (2009). A brief history of human brain mapping. *Trends Neurosci*, *32*(2), 118-26. doi: 10.1016/j.tins.2008.11.001

Rilling, J. K., Glasser, M. F., Jbabdi, S., Andersson, J., & Preuss, T. M. (2011). Continuity, divergence, and the evolution of brain language pathways. *Front Evol Neurosci*, *3*, 11. doi: 10.3389/fnevo.2011.00011

Rosas, H. D., Liu, A. K., Hersch, S., Glessner, M., Ferrante, R. J., Salat, D. H., ... Fischl, B. (2002). Regional and progressive thinning of the cortical ribbon in Huntington's disease. *Neurology*, *58*(5), 695-701.

Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philos Trans R Soc Lond B Biol Sci*, *336*(1278), 367-73. doi: 10.1098/rstb.1992.0070

Roth, W. T., Kopell, B. S., Tinklenberg, J. R., Huntsberger, G. E., & Kraemer, H. C. (1975). Reliability of the contingent negative variation and the auditory evoked potential. *Electroencephalogr Clin Neurophysiol*, *38*(1), 45-50.

Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage*, *52*(3), 1059-69. doi: 10.1016/j.neuroimage.2009.10.003

Saberi, K., & Perrott, D. R. (1999). Cognitive restoration of reversed speech. *Nature*, *398*(6730), 760. doi: 10.1038/19652

Salat, D. H., Buckner, R. L., Snyder, A. Z., Greve, D. N., Desikan, R. S. R., Busa, E., ... Fischl, B. (2004). Thinning of the cerebral cortex in aging. *Cereb Cortex*, *14*(7), 721-30. doi: 10.1093/cercor/bhh032

Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M.-S., ... Weiller, C. (2008). Ventral and dorsal pathways for language. *Proc Natl Acad Sci U S A*, *105*(46), 18035-40. doi: 10.1073/pnas.0805234105

Scarff, C. J., Reynolds, A., Goodyear, B. G., Ponton, C. W., Dort, J. C., & Eggermont, J. J. (2004). Simultaneous 3-T fMRI and high-density recording of human auditory evoked potentials. *Neuroimage*, *23*(3), 1129-42. doi: 10.1016/j.neuroimage.2004.07.035

Schmidt, C. F., Zaehle, T., Meyer, M., Geiser, E., Boesiger, P., & Jancke, L. (2008). Silent and continuous fMRI scanning differentially modulate activation in an auditory language comprehension task. *Hum Brain Mapp*, *29*(1), 46-56. doi: 10.1002/hbm.20372

Schmithorst, V. J., & Holland, S. K. (2004). Event-related fMRI technique for auditory processing with hemodynamics unrelated to acoustic gradient noise. *Magn Reson Med*, *51*(2), 399-402. doi: 10.1002/mrm.10706

Schönwiesner, M., Rübsamen, R., & von Cramon, D. Y. (2005). Hemispheric asymmetry for spectral and temporal processing in the human antero-lateral auditory belt cortex. *Eur J Neurosci*, *22*(6), 1521-8. doi: 10.1111/j.1460-9568.2005.04315.x

Schwartz, J., & Tallal, P. (1980). Rate of acoustic change may underlie hemispheric specialization for speech perception. *Science*, *207*(4437), 1380-1.

Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, *123 Pt 12*, 2400-6.

Ségonne, F., Dale, A. M., Busa, E., Glessner, M., Salat, D., Hahn, H. K., & Fischl, B. (2004). A hybrid approach to the skull stripping problem in MRI. *Neuroimage*, *22*(3), 1060-75. doi: 10.1016/j.neuroimage.2004.03.032

Shah, N. J., Steinhoff, S., Mirzazade, S., Zafiris, O., Grosse-Ruyken, M. L., Jäncke, L., & Zilles, K. (2000). The effect of sequence repeat time on auditory cortex stimulation during phonetic discrimination. *Neuroimage*, *12*(1), 100-8. doi: 10.1006/nimg.2000.0588

Shapleske, J., Rossell, S. L., Woodruff, P. W., & David, A. S. (1999). The planum temporale: a systematic, quantitative review of its structural, functional and clinical significance. *Brain Res Brain Res Rev*, *29*(1), 26-49.

Sigalovsky, I. S., Fischl, B., & Melcher, J. R. (2006). Mapping an intrinsic mr property of gray matter in auditory cortex of living humans: a possible marker for primary cortex and hemispheric differences. *Neuroimage*, *32*(4), 1524-37. doi: 10.1016/j.neuroimage.2006.05.023

Slevc, L. R., Martin, R. C., Hamilton, A. C., & Joanisse, M. F. (2011). Speech perception, rapid temporal processing, and the left hemisphere: a case study of unilateral pure word deafness. *Neuropsychologia*, *49*(2), 216-30. doi: 10.1016/j.neuropsychologia.2010.11.009

Sporns, O., Tononi, G., & Kötter, R. (2005). The human connectome: A structural description of the human brain. *PLoS Comput Biol*, *1*(4), e42. doi: 10.1371/journal.pcbi.0010042

Studdert-Kennedy, M., & Shankweiler, D. (1970). Hemispheric specialization for speech perception. *J Acoust Soc Am*, *48*(2), 579-94.

Talavage, T. M., & Edmister, W. B. (2004). Nonlinearity of fMRI responses in human auditory cortex. *Hum Brain Mapp*, *22*(3), 216-28. doi: 10.1002/hbm.20029

Talavage, T. M., Edmister, W. B., Ledden, P. J., & Weisskoff, R. M. (1999). Quantitative assessment of auditory cortex responses induced by imager acoustic noise. *Hum Brain Mapp*, *7*(2), 79-88.

Tallal, P., Miller, S. L., Bedi, G., Byma, G., Wang, X., Nagarajan, S. S., ... Merzenich, M. M. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science*, *271*(5245), 81-4.

Tamer, G. G., Jr, Luh, W.-M., & Talavage, T. M. (2009). Characterizing response to elemental unit of acoustic imaging noise: an fMRI study. *IEEE Trans Biomed Eng*, *56*(7), 1919-28. doi: 10.1109/TBME.2009.2016573

Thulborn, K. R., Waterton, J. C., Matthews, P. M., & Radda, G. K. (1982). Oxygenation dependence of the transverse relaxation time of water protons in whole blood at high field. *Biochim Biophys Acta*, *714*(2), 265-70.

Tremblay, K. L., & Kraus, N. (2002). Auditory training induces asymmetrical changes in cortical neural activity. *J Speech Lang Hear Res*, *45*(3), 564-72.

Turken, U., & Dronkers, N. F. (2011). The neural architecture of the language comprehension network: converging evidence from lesion and connectivity analyses. *Front Syst Neurosci*, *5*, 1. doi: 10.3389/fnsys.2011.00001

Van Essen, D. C., Drury, H. A., Joshi, S., & Miller, M. I. (1998). Functional and structural mapping of human cerebral cortex: solutions are in the surfaces. *Proc Natl Acad Sci U S A*, *95*(3), 788-95.

Van Lancker Sidtis, D. (2006). Does functional neuroimaging solve the questions of neurolinguistics? *Brain Lang*, *98*(3), 276-90. doi: 10.1016/j.bandl.2006.05.006

Vigneau, M., Beaucousin, V., Hervé, P.-Y., Jobard, G., Petit, L., Crivello, F., ... Tzourio-Mazoyer, N. (2011). What is right-hemisphere contribution to phonological, lexico-semantic, and sentence processing? insights from a meta-analysis. *Neuroimage*, *54*(1), 577-93. doi: 10.1016/j.neuroimage.2010.07.036

Voets, N. L., Hough, M. G., Douaud, G., Matthews, P. M., James, A., Winmill, L., ... Smith, S. (2008). Evidence for abnormalities of cortical development in adolescent-onset schizophrenia. *Neuroimage*, *43*(4), 665-75. doi: 10.1016/j.neuroimage.2008.08.013

Walhovd, K. B., & Fjell, A. M. (2002). One-year test-retest reliability of auditory ERPs in young and old adults. *Int J Psychophysiol*, *46*(1), 29-40.

Walker, K. M. M., Ahmed, B., & Schnupp, J. W. H. (2008). Linking cortical spike pattern codes to auditory perception. *J Cogn Neurosci*, *20*(1), 135-52. doi: 10.1162/jocn.2008.20012

Warrier, C., Wong, P., Penhune, V., Zatorre, R., Parrish, T., Abrams, D., & Kraus, N. (2009). Relating structure to function: Heschl's gyrus and acoustic processing. *J Neurosci*, *29*(1), 61-9. doi: 10.1523/JNEUROSCI.3489-08.2009

Wen, W., Zhu, W., He, Y., Kochan, N. A., Reppermund, S., Slavin, M. J., ... Sachdev, P. (2011). Discrete neuroanatomical networks are associated with specific cognitive abilities in old age. *J Neurosci*, *31*(4), 1204-12. doi: 10.1523/JNEUROSCI.4085-10.2011

Wernicke, C. (1874). *Der aphasische Symptomenkomplex*. Breslau: Max Cohn & Weigert.

Winkler, A. M., Kochunov, P., Blangero, J., Almasy, L., Zilles, K., Fox, P. T., ... Glahn, D. C. (2010). Cortical thickness or grey matter volume? The importance of selecting the phenotype for imaging genetics studies. *Neuroimage*, *53*(3), 1135-46. doi: 10.1016/j.neuroimage.2009.12.028

Wise, R. J., Scott, S. K., Blank, S. C., Mummery, C. J., Murphy, K., & Warburton, E. A. (2001). Separate neural subsystems within 'Wernicke's area'. *Brain*, *124*(Pt 1), 83-95.

Wong, P. C. M., Warrier, C. M., Penhune, V. B., Roy, A. K., Sadehh, A., Parrish, T. B., & Zatorre, R. J. (2008). Volume of left Heschl's gyrus and linguistic pitch learning. *Cereb Cortex*, *18*(4), 828-36. doi: 10.1093/cercor/bhm115

Woods, D. L., Stecker, G. C., Rinne, T., Herron, T. J., Cate, A. D., Yund, E. W., ... Kang, X. (2009). Functional maps of human auditory cortex: effects of acoustic features and attention. *PLoS One*, *4*(4), e5183. doi: 10.1371/journal.pone.0005183

Wright, I. C., McGuire, P. K., Poline, J. B., Travere, J. M., Murray, R. M., Frith, C. D., ... Friston, K. J. (1995). A voxel-based method for the statistical analysis of gray and white matter density applied to schizophrenia. *Neuroimage*, *2*(4), 244-52. doi: 10.1006/nimg.1995.1032

Xiang, H.-D., Fonteijn, H. M., Norris, D. G., & Hagoort, P. (2010). Topographical functional connectivity pattern in the perisylvian language networks. *Cereb Cortex*, *20*(3), 549-60. doi: 10.1093/cercor/bhp119

Xu, Y., Gandour, J., Talavage, T., Wong, D., Dzemidzic, M., Tong, Y., ... Lowe, M. (2006). Activation of the left planum temporale in pitch processing is shaped by language experience. *Hum Brain Mapp*, *27*(2), 173-83. doi: 10.1002/hbm.20176

Yvert, B., Crouzeix, A., Bertrand, O., Seither-Preisler, A., & Pantev, C. (2001). Multiple supratemporal sources of magnetic and electric auditory evoked middle latency components in humans. *Cereb Cortex*, *11*(5), 411-23.

Zaehle, T., Jancke, L., Herrmann, C. S., & Meyer, M. (2009). Pre-attentive spectro-temporal feature processing in the human auditory system. *Brain Topogr*, *22*(2), 97-108. doi: 10.1007/s10548-009-0085-6

Zaehle, T., Jancke, L., & Meyer, M. (2007). Electrical brain imaging evidences left auditory cortex involvement in speech and non-speech discrimination based on temporal features. *Behav Brain Funct*, *3*, 63. doi: 10.1186/1744-9081-3-63

Zaehle, T., Schmidt, C. F., Meyer, M., Baumann, S., Baltes, C., Boesiger, P., & Jancke, L. (2007). Comparison of "silent" clustered and sparse temporal fMRI acquisitions in tonal and speech perception tasks. *Neuroimage*, *37*(4), 1195-204. doi: 10.1016/j.neuroimage.2007.04.073

Zaehle, T., Wüstenberg, T., Meyer, M., & Jäncke, L. (2004). Evidence for rapid auditory perception as the foundation of speech processing: a sparse temporal sampling fMRI study. *Eur J Neurosci*, *20*(9), 2447-56. doi: 10.1111/j.1460-9568.2004.03687.x

Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cereb Cortex*, *11*(10), 946-53.

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends Cogn Sci*, *6*(1), 37-46.

Zatorre, R. J., Fields, R. D., & Johansen-Berg, H. (2012). Plasticity in gray and white: neuroimaging changes in brain structure during learning. *Nat Neurosci*, *15*(4), 528-36. doi: 10.1038/nn.3045

Zhang, L., Shu, H., Zhou, F., Wang, X., & Li, P. (2010). Common and distinct neural substrates for the perception of speech rhythm and intonation. *Hum Brain Mapp*, *31*(7), 1106-16. doi: 10.1002/hbm.20922

Zurif, E., & Mendelsohn, M. (1972). Hemispheric specialization for the perception of speech sounds: The influence of intonation and structure. *Attention, Perception, & Psychophysics*, *11*, 329-332. (10.3758/BF03206262)

A. Supplementary Figures

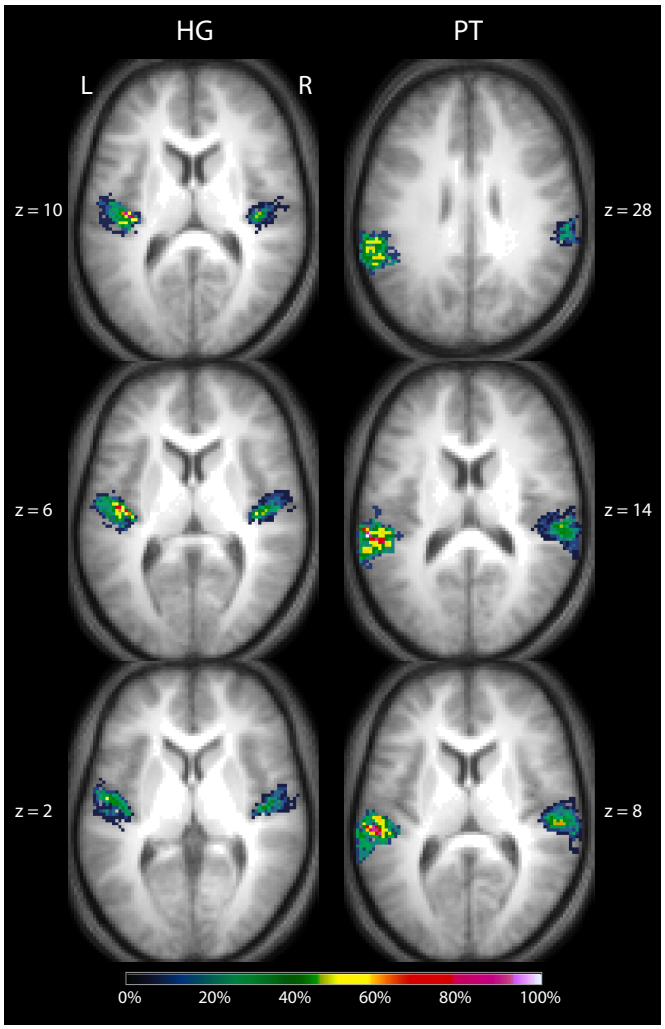

Figure A.1.: Structural Overlap maps of the individual ROIs. HG: Heschl's gyrus; PT: planum temporale. MNI space. Neurological convention.

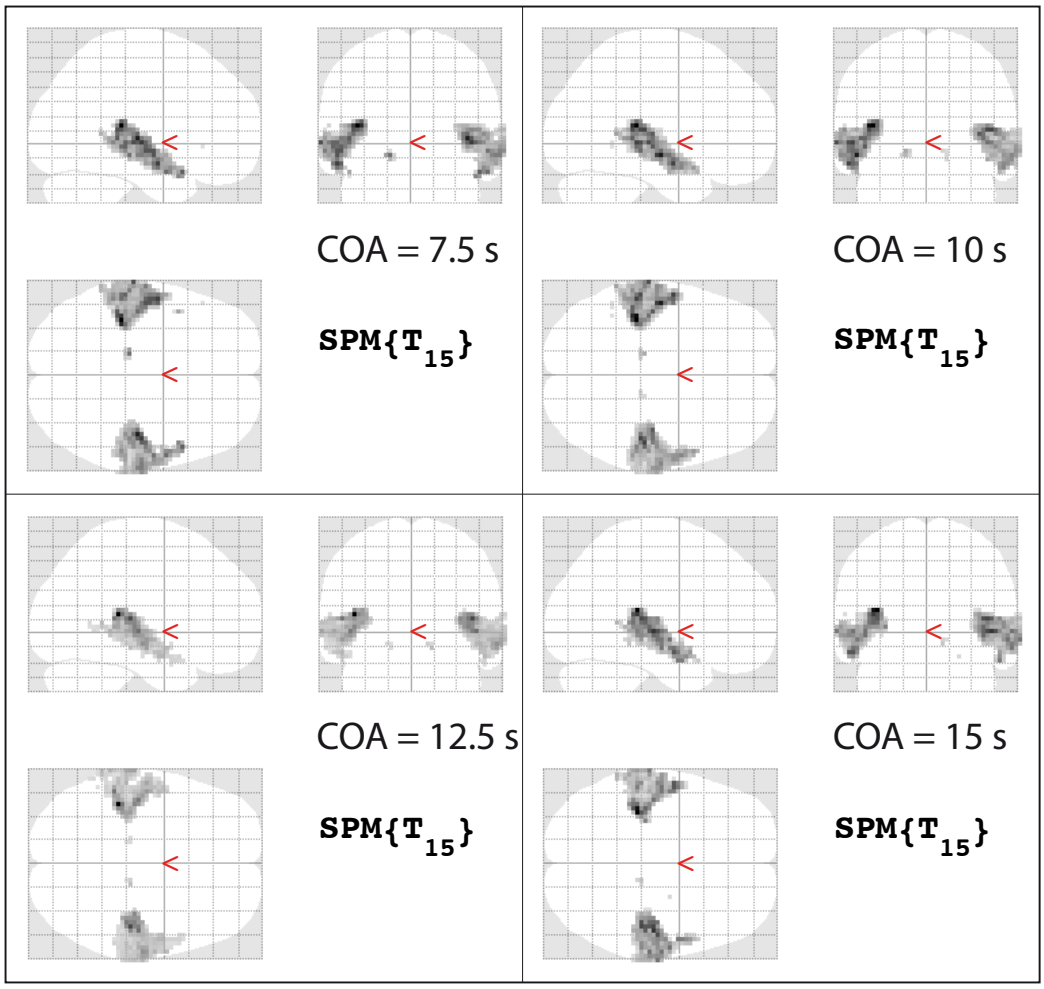

Figure A.2.: Glassbrain projections of the T-contrast for all auditory events vs. empty baseline at all COA settings; (FWE, $p < .05$, $T > 6.8$, $k = 0$). MNI space. Neurological convention.

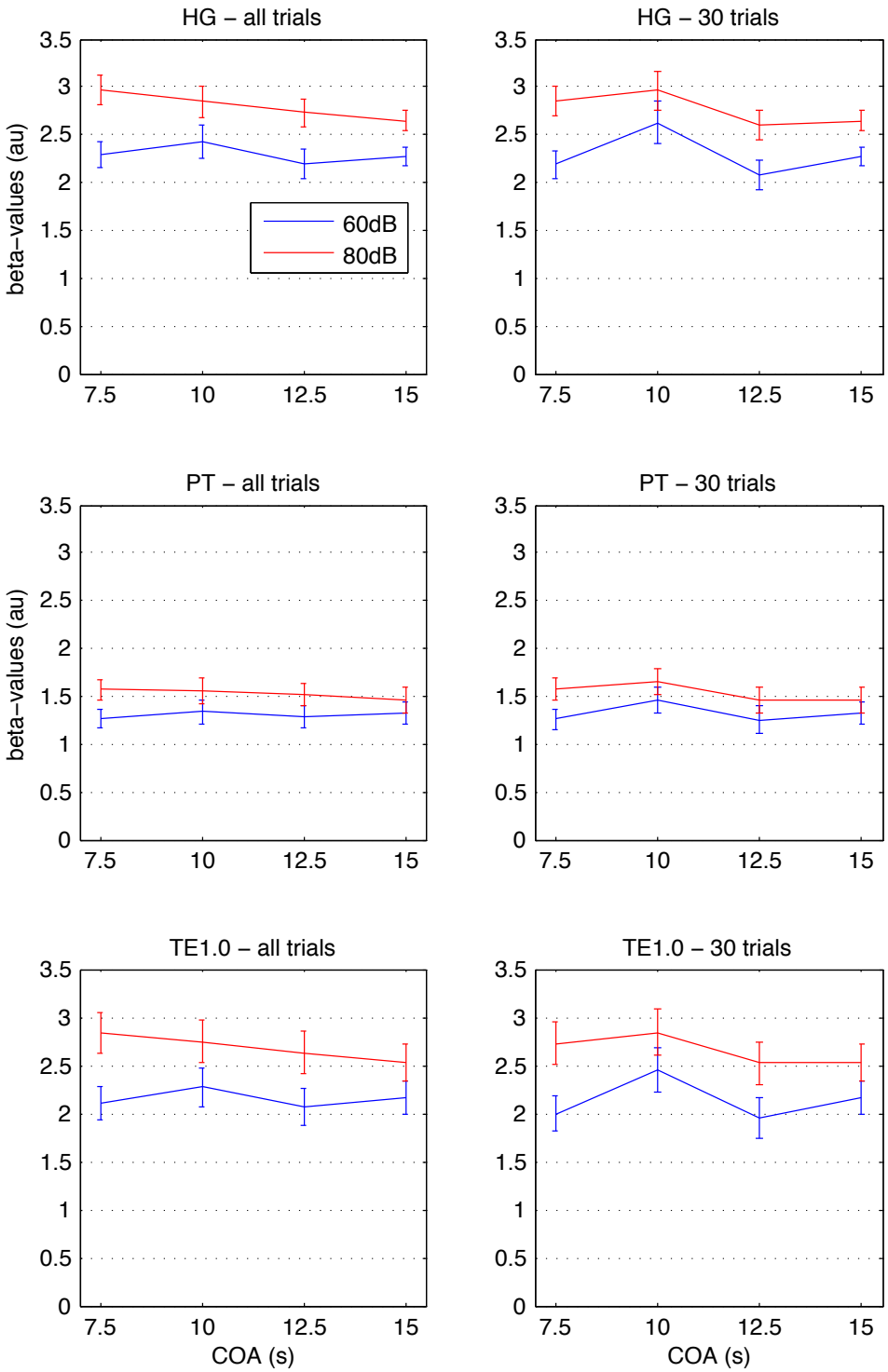

Figure A.3.: Mean beta-values (60 vs. rest and 80 dB vs. rest) for each of the four COA settings in all three ROIs (± SEM).

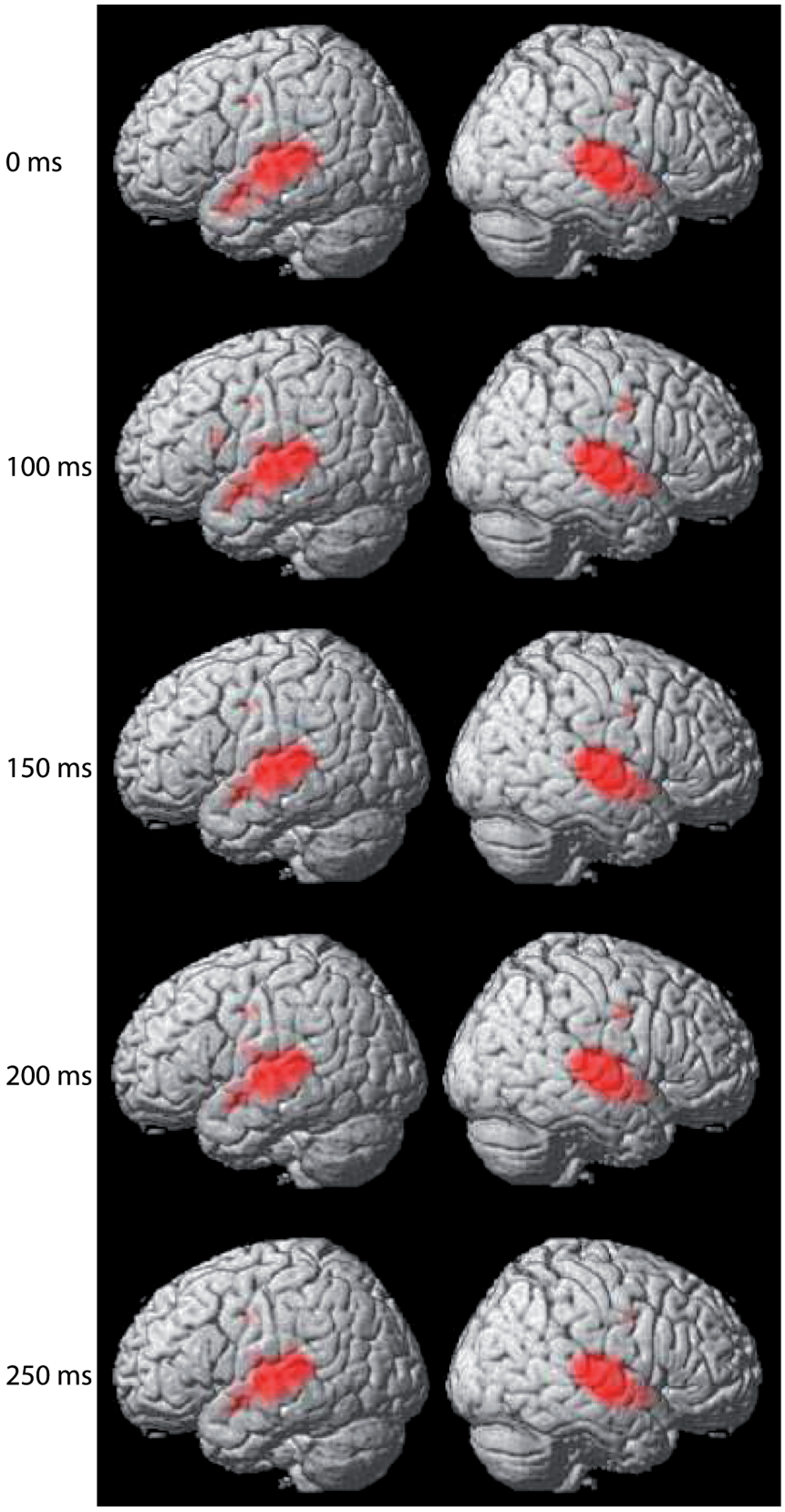

Figure A.4.: Rendering of the group contrasts conditions vs. rest (p(FWE) $< .05$).

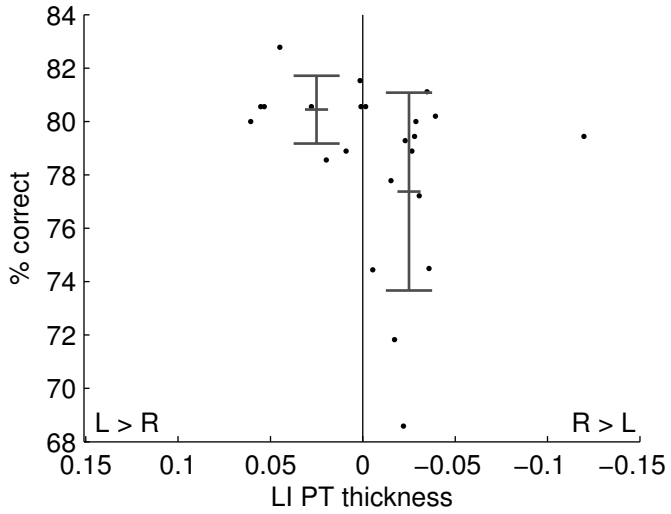

Figure A.5.: Scatterplot of lateralization index of planum temporale (LI PT) thickness by percent correct answers. Bars indicate M and SD of the PT groups.

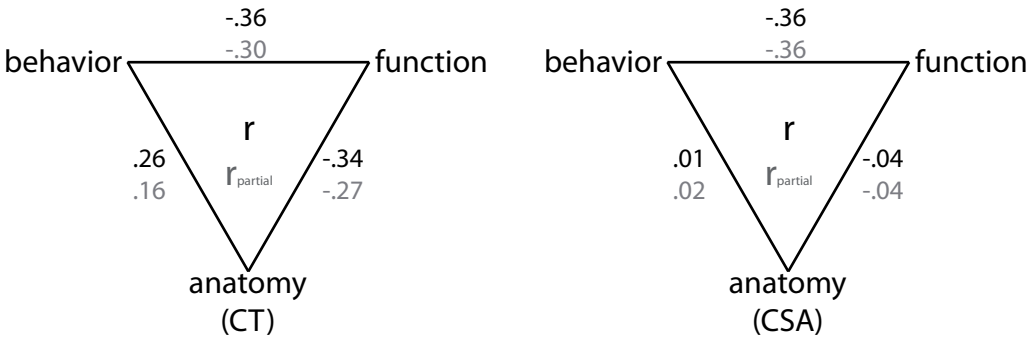

Figure A.6.: Correlations (black) and partial correlations (grey) between behavior (mean percent correct), function (functional lateralization in PT, mean of all five conditions) and anatomy (PT asymmetry) for cortical thickness (CT, left) and cortical surface area (CSA, right), respectively. No correlation reached statistical significance.